

N 64 81342

Code None

164p

Volume III of III

Unclassified Title

R-273

INTERPLANETARY NAVIGATION
SYSTEM STUDY

April, 1960

by

J. H. Laning, Jr.

J. M. Dahlen

R. Battin

R. J. Magee

P. Bowditch

K. Nordtvedt

M. B. Trageser

R. A. Scholten

R. L. Alonso

H. H. Seward

W. E. Toth

INSTRUMENTATION
LABORATORY •

MASSACHUSETTS

INSTITUTE

OF

TECHNOLOGY

Cambridge

39,

Mass.

DOUGLAS, LEO
FBI

OCT 20 3 53 PM 1961

TEM
CITECO

CASE FILE COPY

N 64 81342 ²⁴

Volume III of III
Unclassified Title

Code None

R-273

1 INTERPLANETARY NAVIGATION SYSTEM STUDY

by

J. H. Laning, Jr.,

J. M. Dahlen

R. Battin,

R. J. Magee

P. Bowditch,

K. Nordtvedt

M. B. Trageser,

R. A. Scholten

R. L. Alonso *et al*

H. H. Seward

W. E. Toth

also

1st 5 pages

auths. and index

all 11 (NASA support)

April, 1960

164 p refs

(2)
1486984
INSTRUMENTATION LABORATORY
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
CAMBRIDGE 39, MASSACHUSETTS

Approved by:

Roger B. Woodbury

Associate Director

Contract AF 04(647)-303; (NASA Contract NASw-130)
Contract NORD-17366)
(NASA CR-55920;
1 12-273)
Index only CN-83977, Vol. 3

ACKNOWLEDGMENT

With the exception of Appendix E, this report was prepared under DSR Project 55-171, sponsored by the National Aeronautics and Space Administration through contract NASw -130.

Appendix E was prepared under DSR Project 52-156 sponsored by the Ballistic Missile Division of the Air Research and Development Command through USAF Contract AF 04(647)-303 and Project 53-138, Division of Sponsored Research, Massachusetts Institute of Technology, sponsored by the Bureau of Ordnance, Department of the Navy, under Contract NOrd17366.

The cost of reproducing Appendix E here was borne by DSR Project 55-171.

TABLE OF CONTENTS

	Volume
Chapter 1 General Considerations and Summary	I
Chapter 2 Orbit Studies	I
Chapter 3 Navigation Studies	I
Chapter 4 A Centaur Interplanetary Spacecraft Guidance and Control System	I
Chapter 5 Additional Subsystem Studies	II
Chapter 6 Environmental Problems, Tests, Facilities and Program	II
Chapter 7 Sublimating-Solid Torquing Jet	II
Chapter 8 Proposed Advanced Development Program Schedules, and Required Funding	II
Appendix A Entry into Planetary Atmosphere	III
Appendix B Computational Procedures for the Navigational Fix	III
Appendix C Variable-Time-of-Arrival Naviga- tion Theory	III
Appendix D Approximate Position and Velocity of the Moon	III
Appendix E Design Principles for a General Control Computer	III
Supplement (separately bound)	

APPENDIX A
ENTRY INTO PLANETARY ATMOSPHERES

by
Roger A. Scholten

TABLE OF CONTENTS

Appendix		Page
A	Entry in Planetary Atmosphere	9
B	Computational Procedures for the Navigational Fix	25
C	Variable-Time-of-Arrival Navigation Theory	35
D	Approximate Position and Velocity of the Moon	53
E	Design Principles for a General Control Computer	65
	Chapter 1 - General Organization of Storage	75
	Chapter 2 - Execution of Computer Instructions	97
	Chapter 3 - Input-Output Systems	131
	Chapter 4 - General Remarks	161

LIST OF ILLUSTRATIONS

Figure		
A-1	Motion of vehicle in vicinity of destination planet	10
A-2	Maximum drag deceleration during entry into Earth's atmosphere $W/C_D A$ between 50 and 100 lbs/ft ²	15
A-3	Range angle traveled during entry into Earth's atmosphere for $W/C_D A$ between 50 and 100 lbs/ft ²	16
A-4	Altitude at maximum deceleration for various conditions of entry into Earth's atmosphere	17
A-5	Maximum drag deceleration during entry into Mars' atmosphere $W/C_D A$ between 50 and 100 lbs/ft ²	18
A-6	Range angle traveled during entry into Mars' atmosphere for $W/C_D A$ between 50 and 100 lbs/ft ²	19
A-7	Altitude at maximum deceleration for various conditions of entry into Mars' atmosphere	20
A-8	Maximum drag deceleration during entry into Venus' atmosphere for $W/C_D A$ between 50 and 100 lbs/ft ²	21
A-9	Range angle traveled during entry into Venus' atmosphere for $W/C_D A$ between 50 and 100 lbs/ft ²	22
A-10	Altitude at maximum deceleration for various conditions of entry into Venus' atmosphere	23
D-1	Geometry of moon's position relative to the Earth	59
1-1	Word core rope	76
1-2	Selection of a core in a core rope	78
1-3	Timing of inhibiting, set and clear currents	79
1-4	Sensing wires on a core rope	80

Figure		Page
1-5	Symbolic erasable register	82
1-6	Erasable register	84
1-7	Erasable storage system	86
1-8	System of ten erasable storage registers	88
1-9	Erasable storage selection system	91
1-10	Complete memory system	92
2-1	Over-all organization of the representative computer . .	98
2-2	Sequence generator	102
2-3	Control pulse generation by sequence generator	104
2-4	Representative central register system.	108
2-5	Average number of cycles for addition	127
3-1	Priority circuit	133
3-2	Priority circuit system	135
3-3	Input circuits for priority circuit system	136
3-4	Unit M for input to priority circuit	141
3-5	Circuitry for state input	143
3-6	Ideal counter register	146
3-7	System for counter inputs	147
3-8	Counter inputs -- system for dealing with positive and negative increments	151
3-9	Output switch circuit	153
3-10	Gating of a pulse generator by an output switch	154
3-11	Generator of several pulse rates. Pulse rate $Q = a_1 R + a_2 \frac{R}{2} + a_3 \frac{R}{4} + a_4 \frac{R}{8} + a_5 \frac{R}{16}$, where $a_1 = 0$ or 1 , and R is some pulse rate determined by the clock. The contents of an output register determine which a are 0 and which are 1	156
3-12	Details of an output rate register	157
3-13	Pulse rate generation and examples	159
4-1	Sixteen word erasable storage system. (Clearing tran- sistors at center. Write transistor-core combinations at edges. Write lines are under registers.)	163
4-2	Physical characteristics of components	165

APPENDIX A

ENTRY INTO PLANETARY ATMOSPHERES

The equations of motion of a vehicle moving in the vicinity of a planet have been developed in a form useful for interplanetary landings or passes. ⁽¹⁾

The following extends this material by taking the planetary atmosphere into consideration. The procedure used was to compute conditions just prior to significant atmospheric effects, using equations developed in Ref 1; taking these as initial conditions the rest of the trajectory was obtained by numerical integration.

Impacts on Venus, Earth, and Mars were considered for various relative velocities and aiming points, as well as drag characteristics.

Definitions for the quantities illustrated in Fig. A-1, which was taken from Ref 1, are:

- V_R the velocity of the vehicle relative to the planet, if the planet had no gravitational field (the relative velocity at infinity).
- r_a the distance the vehicle would pass from the center of the planet if gravitational and atmospheric effects did not exist.
- ψ_0 the angle the velocity vector makes with the local vertical at the start of the numerical integration.
- ϕ the angle traveled around the planet by the vehicle.
 $\phi_0 = \phi$ at the start of numerical integration. ϕ is zero

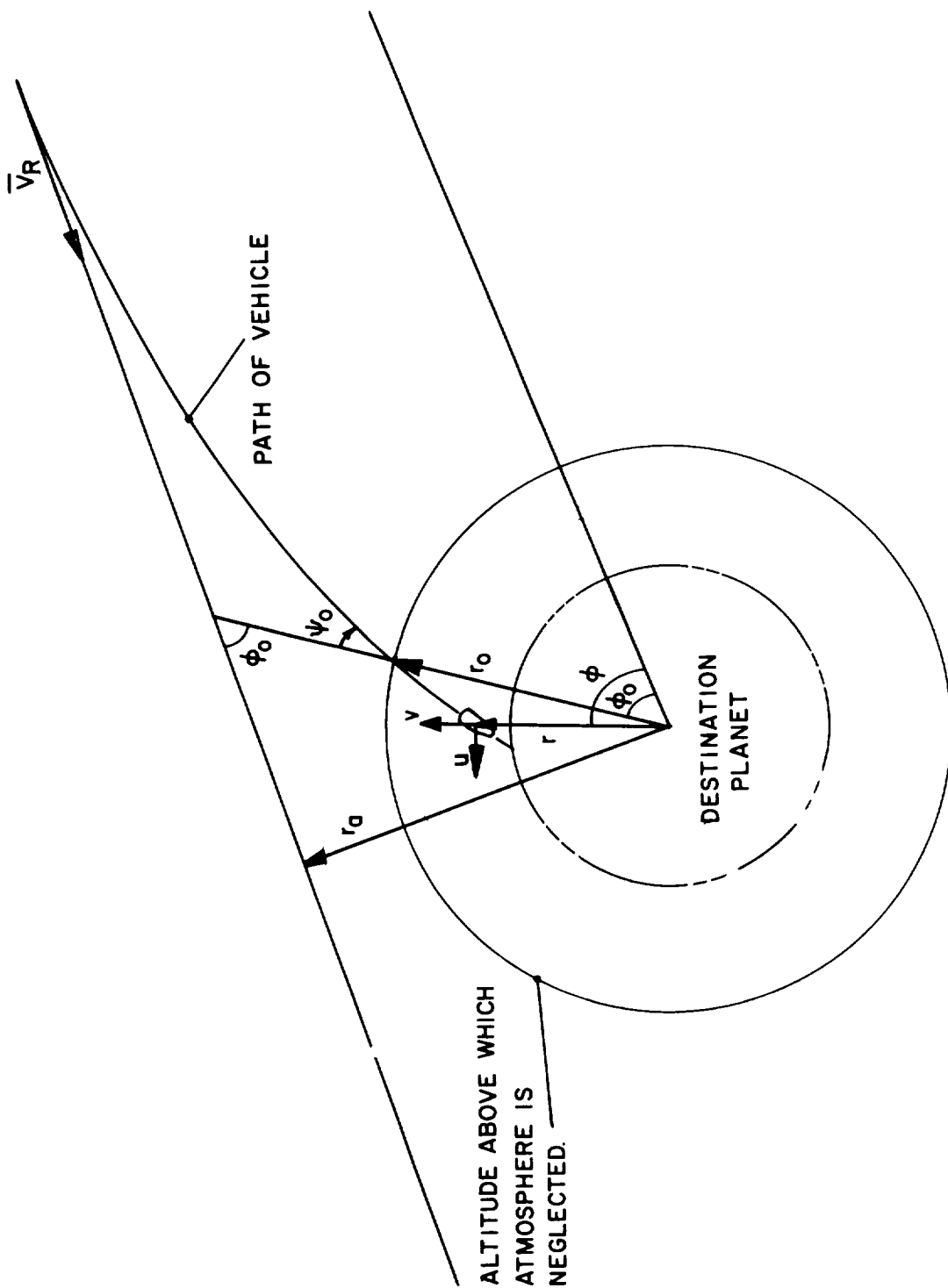


Fig. A-1 Motion of vehicle in vicinity of destination planet.

for a straight-in approach with the same vector velocity, V_R , at infinity.

- r the distance from the vehicle to the center of the planet.
- r_o the radius of the planet plus the height of the atmosphere.
- u the vehicle's horizontal velocity.
- v the vehicle's vertical velocity.

The following procedure was used to determine the vehicle's velocity and position at the start of re-entry.

1. Various reasonable values for the relative velocity (V_R) were selected.
2. For each value of V_R a trajectory was generated using various values of ϕ_o from zero until entry no longer occurred.
3. The distance r_a was computed from (see Ref 1 for derivation):

$$r_a^2 - r_a r_o \sin \phi_o - \frac{\mu_p}{V_R^2} r_o (1 - \cos \phi_o) = 0,$$

where μ_p is the universal gravitation constant times the mass of the planet under consideration.

4. The entry velocity direction was computed from (see Ref 1 for derivation):

$$\tan \psi_o = \frac{r_a^2}{r_o \left(r_a \cos \phi_o + \frac{\mu_p}{V_R^2} \sin \phi_o \right)}.$$

5. The magnitude of the velocity at the beginning of re-entry (V_o) was computed from energy considerations:

$$V_o^2 = V_R^2 + 2 \frac{\mu_p}{r_o}.$$

Using the initial conditions obtained above, the position and velocity of the vehicle as a function of time were found by numerical integration of the following differential equations.

$$\frac{dr}{dt} = v,$$

$$\frac{d\phi}{dt} = \frac{u}{r},$$

$$\frac{du}{dt} = -\frac{uv}{r} - \frac{D_{\text{dec}} u}{(u^2 + v^2)^{1/2}},$$

$$\frac{dv}{dt} = -G - \frac{D_{\text{dec}} v}{(u^2 + v^2)^{1/2}} + u^2 r,$$

where

$$D_{\text{dec}} = \frac{1}{2} \rho_o e^{-\beta h} (u^2 + v^2) \frac{C_D A}{W},$$

the deceleration in ft/second² due to drag forces.

The constants used in these equations are given in Table A-1.

Fig. A-2, A-5, and A-8 show the maximum drag deceleration experienced during re-entry under various conditions of undisturbed passing distance, r_a , and magnitude of relative velocity at infinity, V_R . These curves are useful in determining how far from the planet's center to aim in order to stay within the permissible "g" loading and still be assured of landing.

Fig. A-3, A-6, and A-9 show the range angle traveled by the vehicle around the planet before hitting the surface. This data is useful in predicting landing points and miss coefficients.

Fig. A-4, A-7, and A-10 show the altitude at which maximum deceleration occurs. It can be seen that at steep entry angles the maximum drag occurs at the same altitude for vehi-

cles with the same $W/C_D A$, independent of velocity. However, considerable change in $W/C_D A$ can occur without appreciably affecting the value of maximum drag, which is very sensitive to changes in velocity.

TABLE A-1

ASSUMED VALUES FOR PHYSICAL CHARACTERISTICS

	Venus	Earth	Mars
Radius ⁽²⁾ in terms of Earth's radius of 2.09×10^7 ft.	0.973	1.00	0.532
Mass ⁽²⁾ in terms of Earth's mass	0.820	1.00	0.108
Surface level atmospheric density ⁽³⁾ - lb - ft ⁻³ (ρ_0)	1.00	0.0765	0.0062
Beta ⁽³⁾ used in computing atmospheric density ft ⁻¹	4.88×10^{-5}	4.15×10^{-5}	1.15×10^{-5}
Altitude above which atmosphere is neglected - ft	400,000	300,000	700,000
$W/C_D A$ weight/drag coefficient times area - lb - ft ⁻²	50 and 100	50 and 100	50 and 100
Lift to mass ratio	0	0	0
The universal gravitation constant times the mass of the earth = 14.081476×10^{15} ft ³ -sec ⁻²			

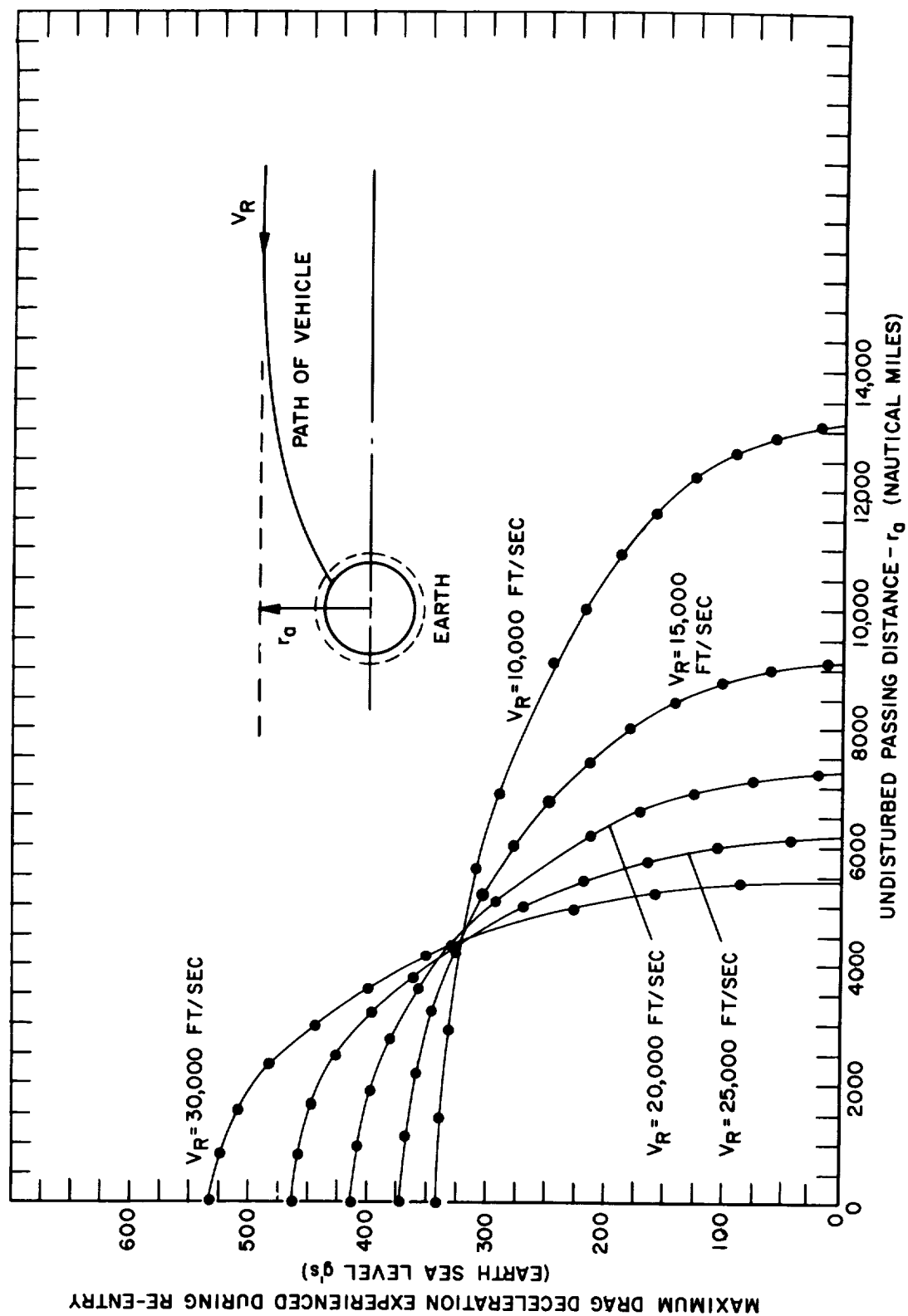


Fig. A-2 Maximum drag deceleration during entry into Earth's atmosphere $W/C_D A$ between 50 and 100 lbs/ft².

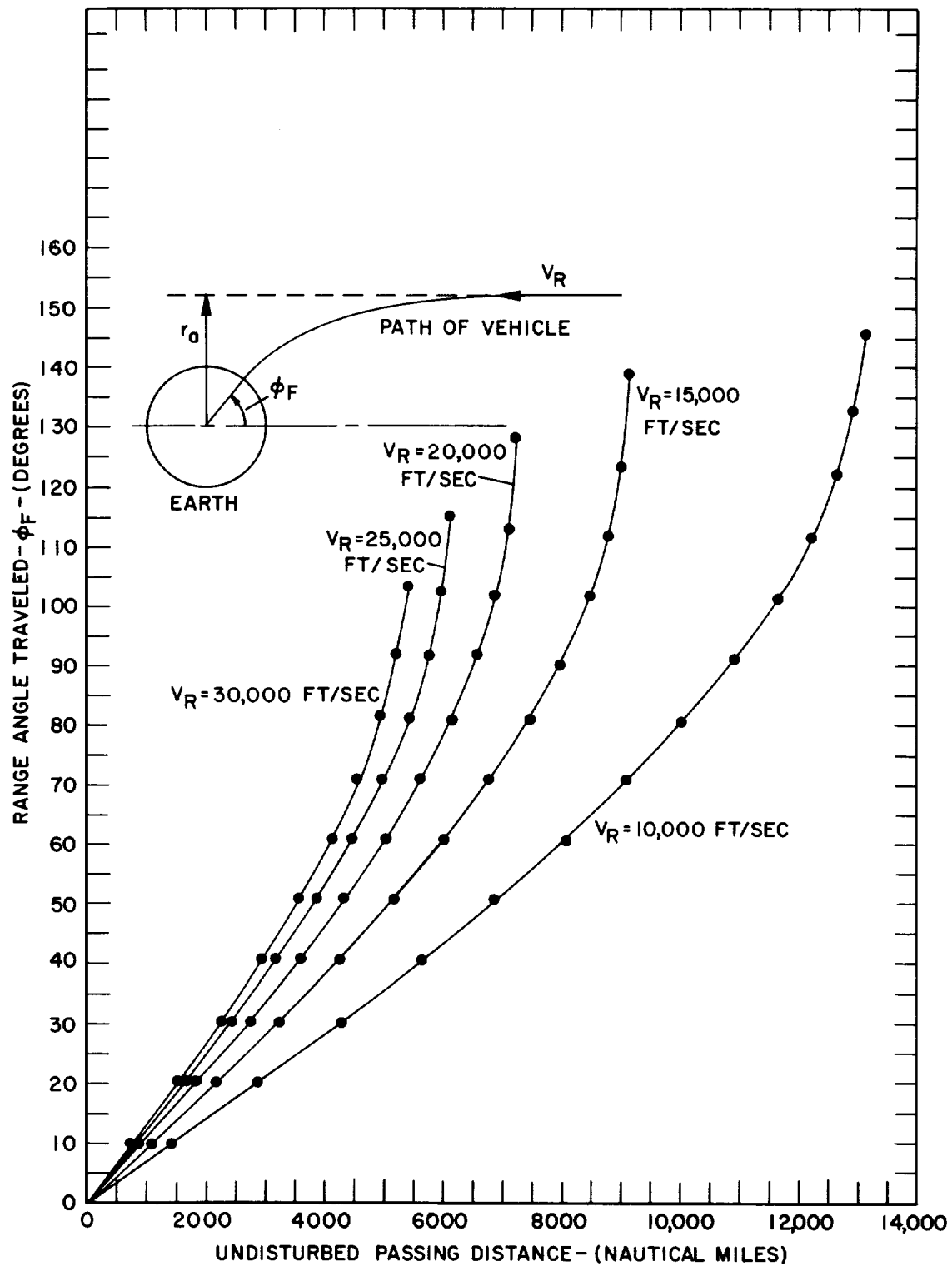


Fig. A-3 Range angle traveled during entry into Earth's atmosphere for $W/C_D A$ between 50 and 100 lbs/ft².

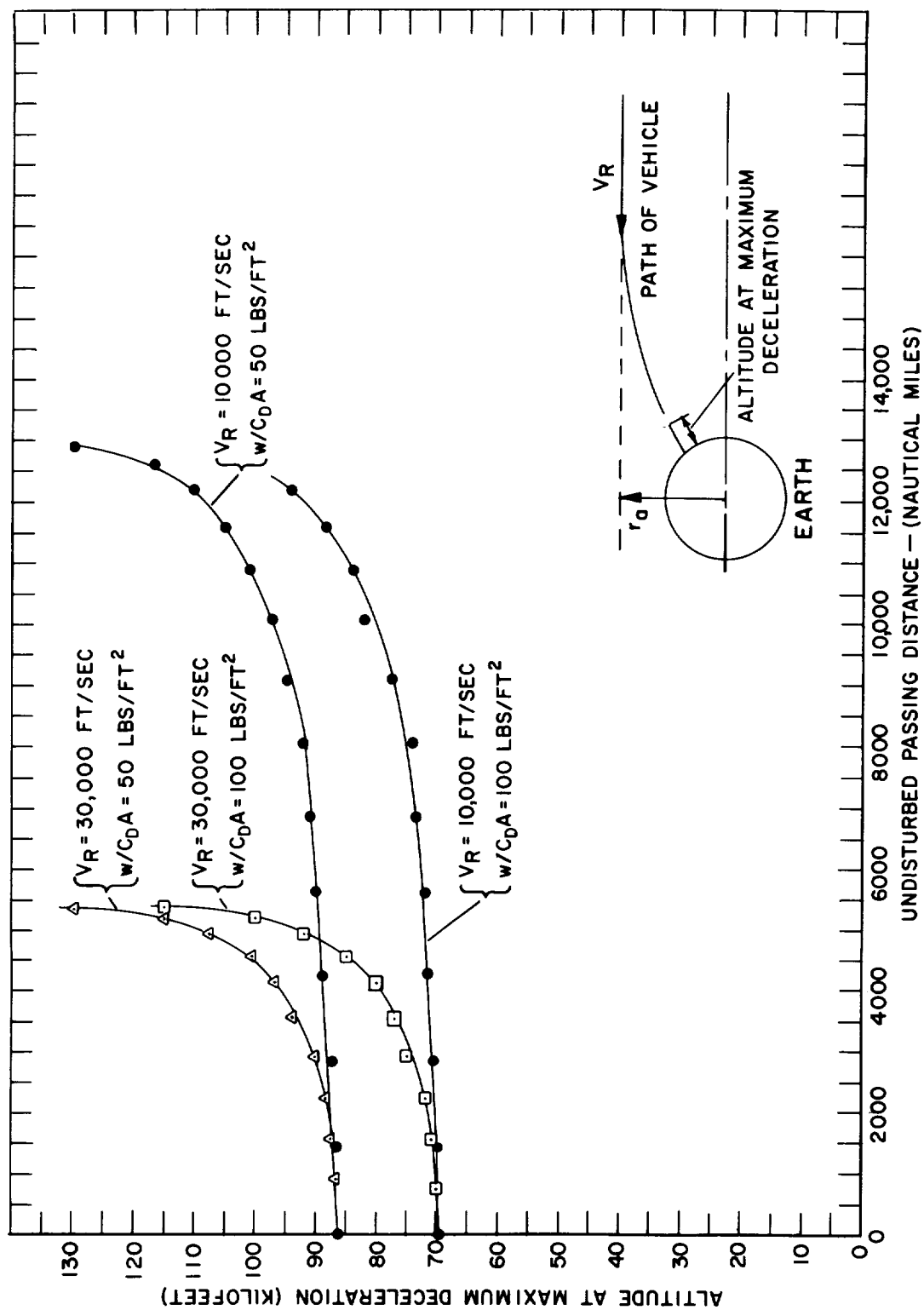


Fig. A-4 Altitude at maximum deceleration for various conditions of entry into Earth's atmosphere.

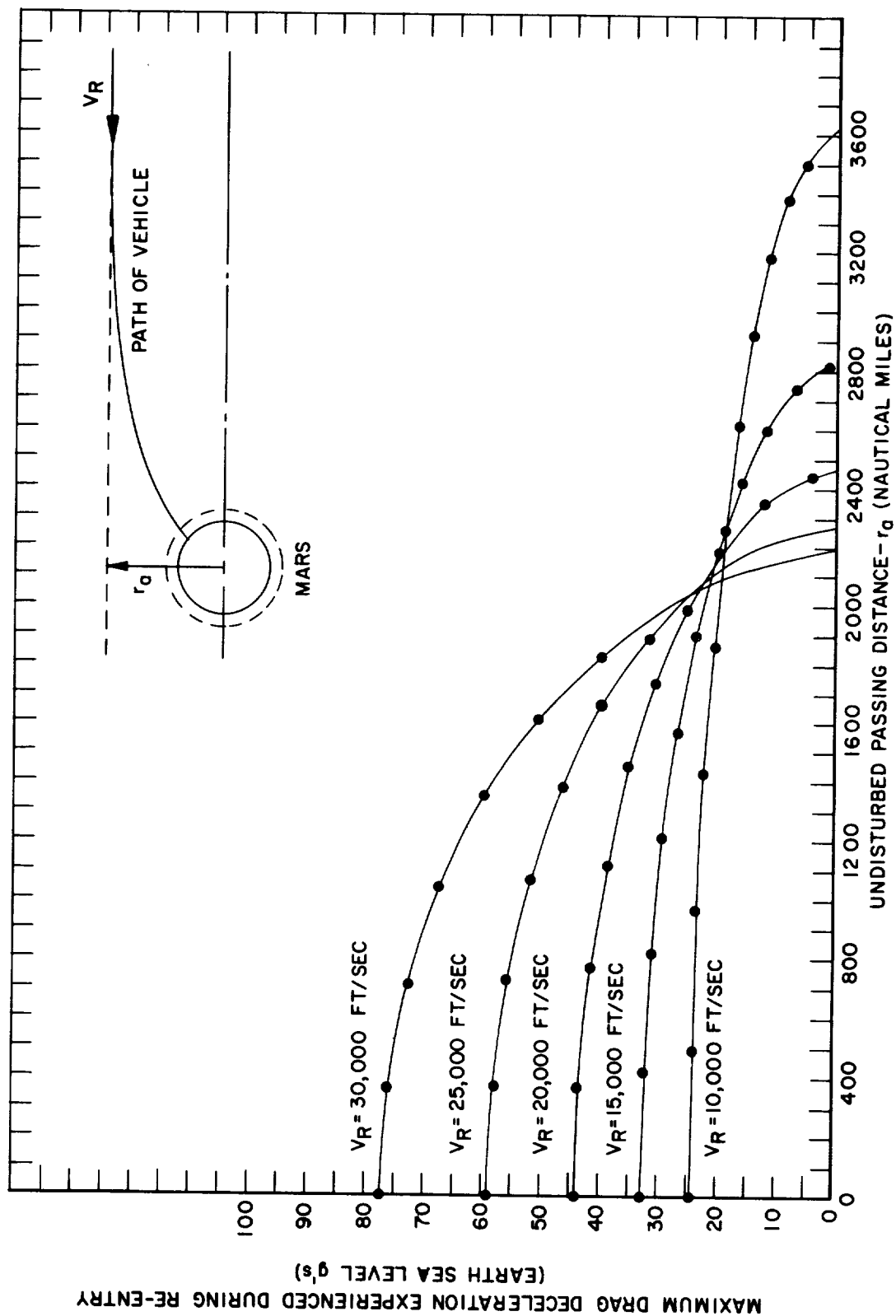


Fig. A-5 Maximum drag deceleration during entry into Mars' atmosphere $W/C_D A$ between 50 and 100 lbs/ft^2 .

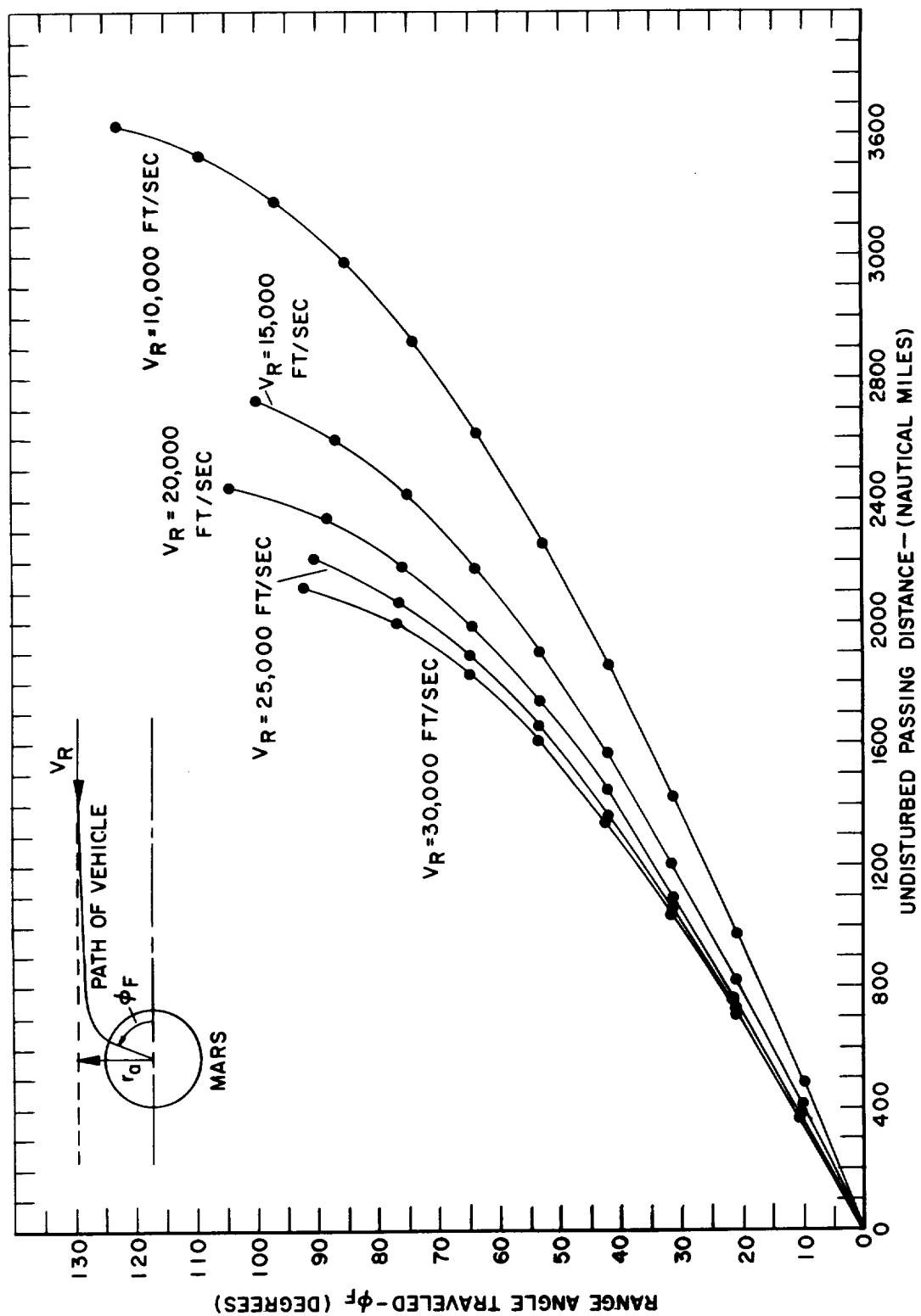


Fig. A-6 Range angle traveled during entry into Mars' atmosphere for W/C_D between 50 and 100 lbs/ft².

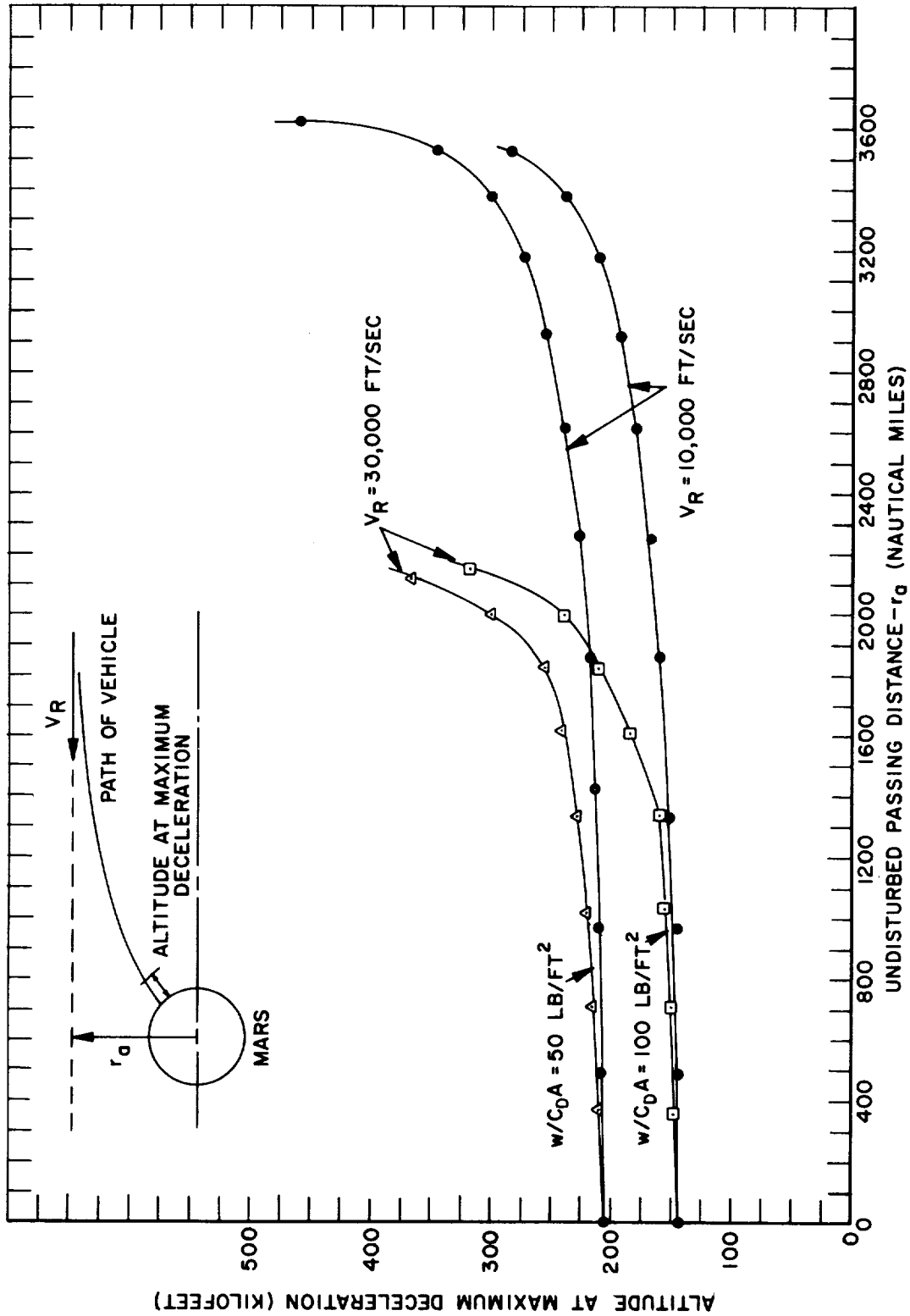


Fig. A-7 Altitude at maximum deceleration for various conditions of entry into Mars' atmosphere.

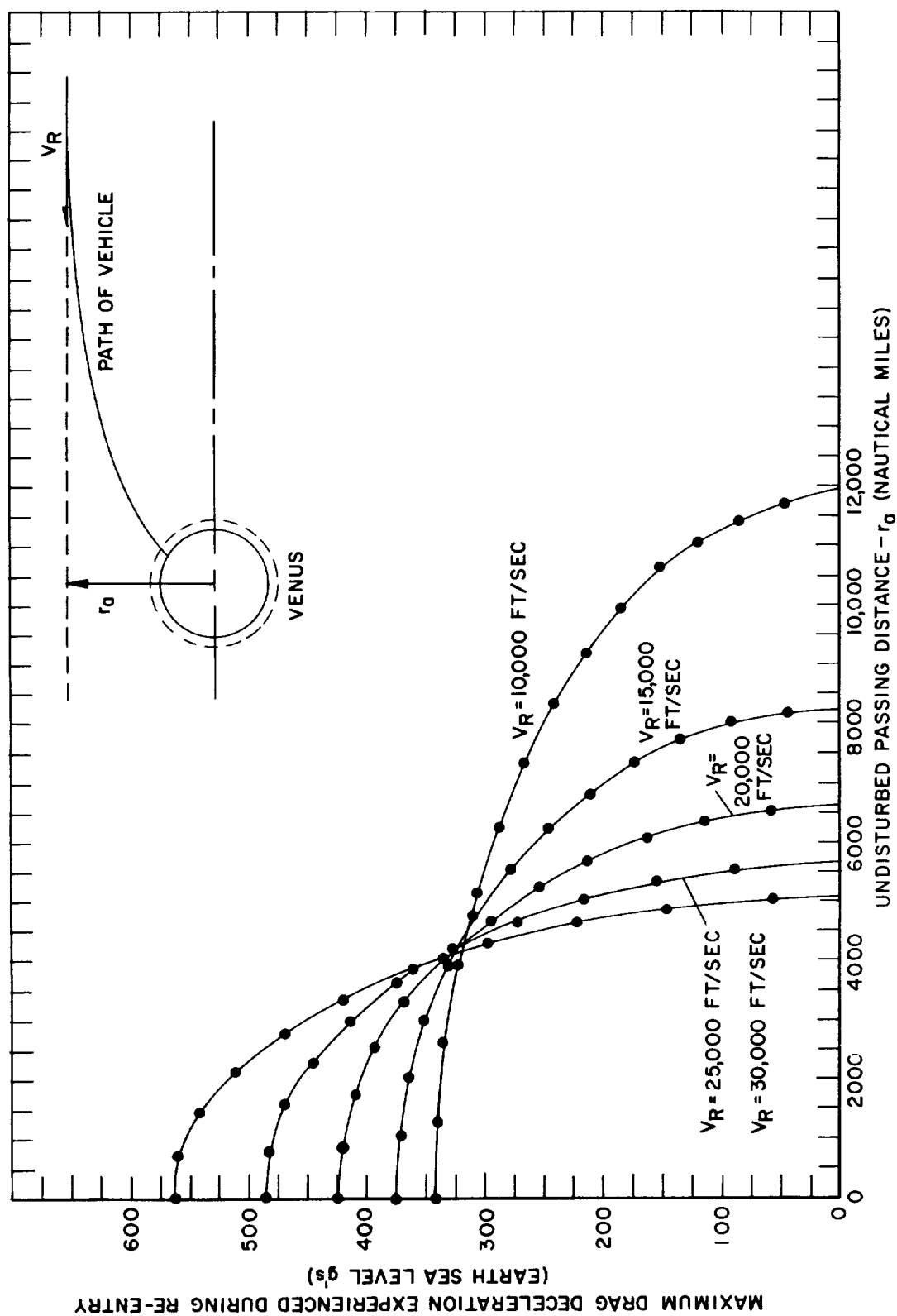


Fig. A-2 Maximum drag deceleration during entry into Venus! atmosphere for W/C_D between 50 and 100 lbs/ft².

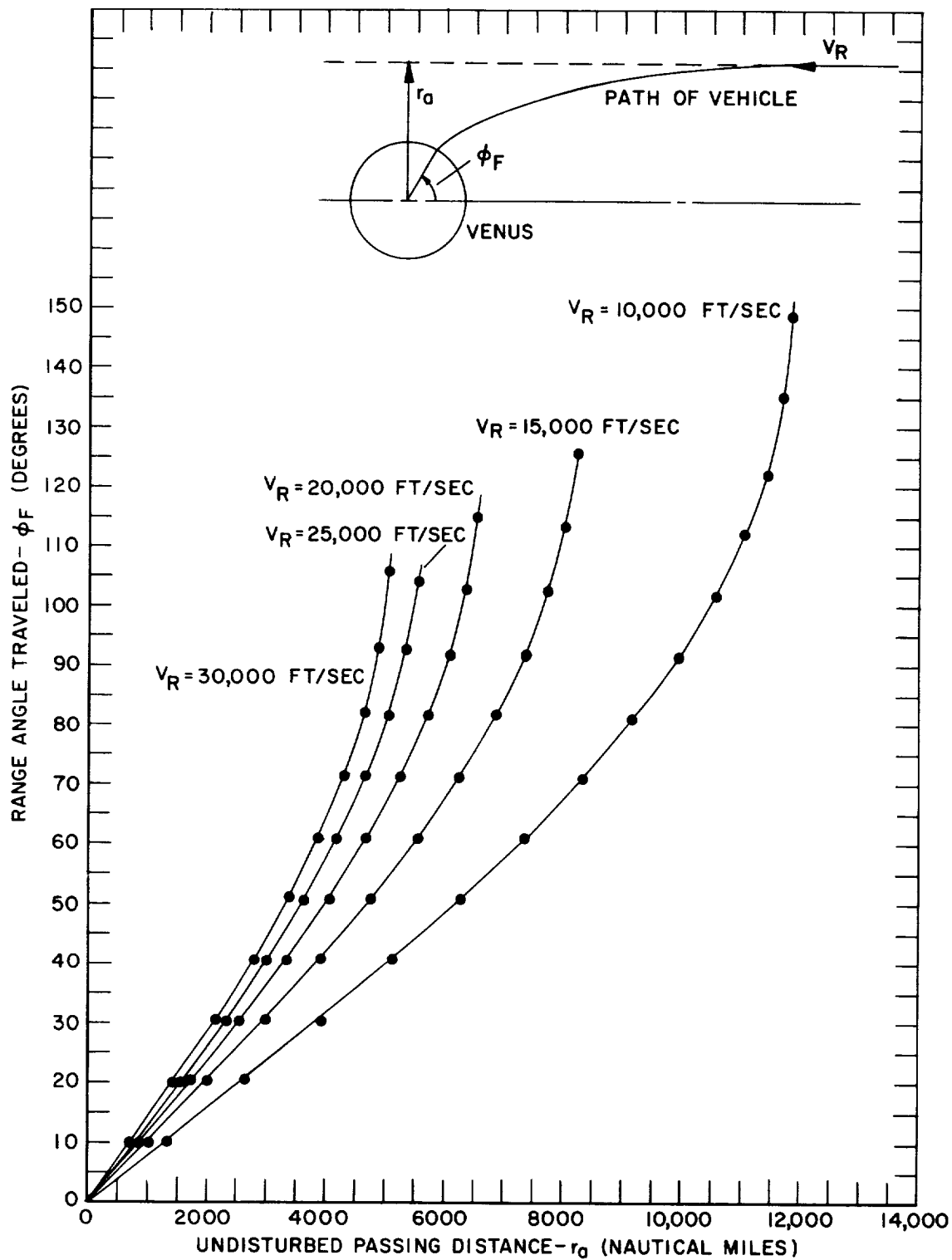


Fig. A-9 Range angle traveled during entry into Venus' atmosphere for $W/C_D A$ between 50 and 100 lbs/ft².

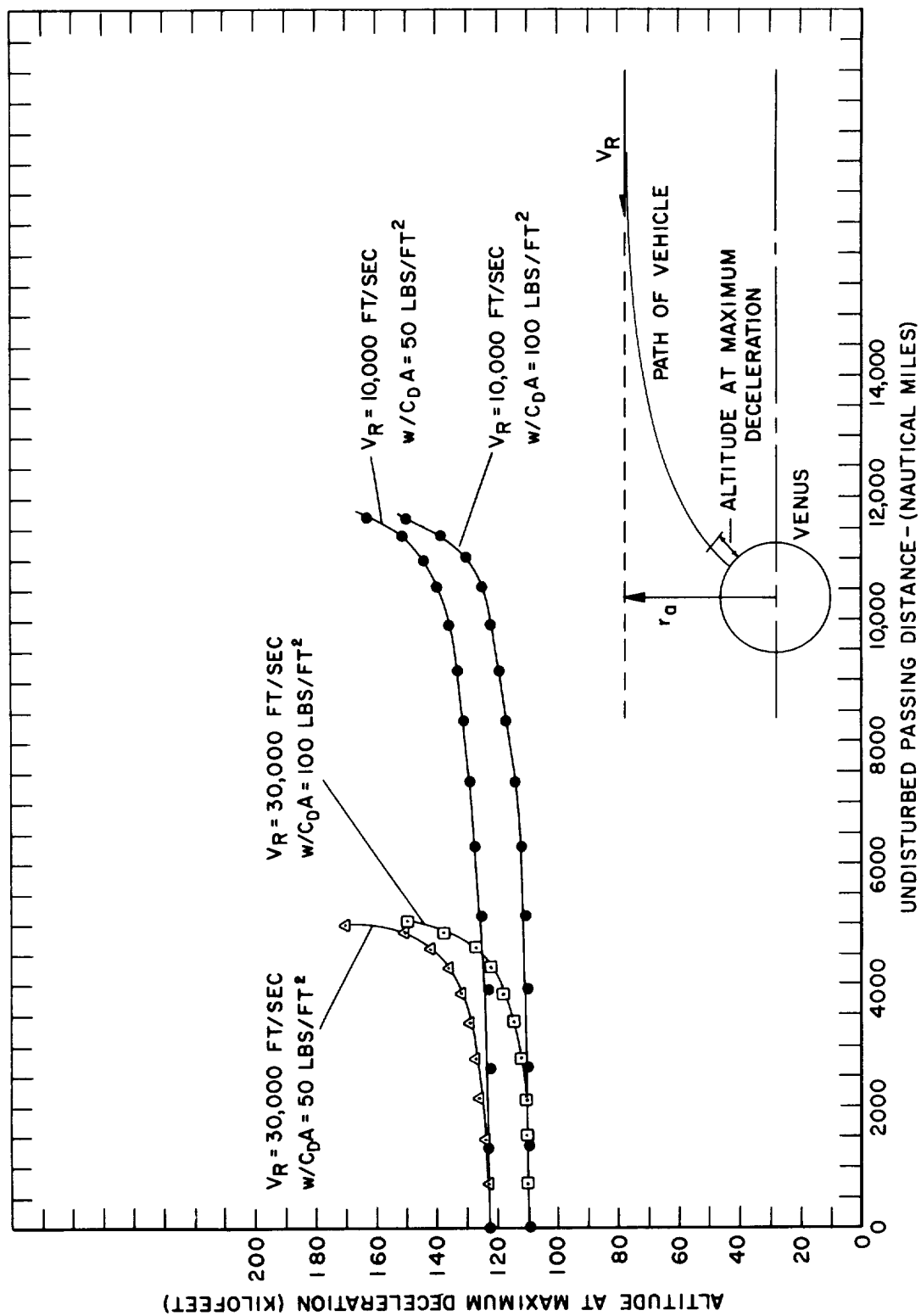


Fig. A-10 Altitude at maximum deceleration for various conditions of entry into Venus' atmosphere.

REFERENCES

1. Battin, Richard, A Recoverable Interplanetary Space Probe, MIT Instrumentation Laboratory Report R-235, Appendix F, July 1959.
2. Russell, Dugan and Stewart, Astronomy I - The Solar System, Ginn and Company, 1945.
3. Gazley, C. , Jr. , Deceleration and Heating of a Body entering A Planetary Atmosphere from Space, Vistas in Astronautics, 1958.

APPENDIX B

COMPUTATIONAL PROCEDURES FOR THE
NAVIGATIONAL FIX

by

Richard H. Battin

TABLE OF CONTENTS

	Page
Appendix B	27
Computational Procedures for the Navigational Fix	

APPENDIX B

COMPUTATIONAL PROCEDURES FOR THE NAVIGATIONAL FIX

The process involved in determining spaceship position by means of a celestial fix consists of a sequence of measurements of the angles between selected pairs of celestial bodies and the angular diameter of a nearby planet when appropriate, together with the recording of time as indicated by the spaceship clock. As a result of these observations a determination of the spaceship position, as well as a correction to the clock reading, may be made. Since the development of this theory, which has been described in R-235, an alternate calculational procedure has been developed which overcomes some of the numerical difficulties inherent in the numerical inversion of a matrix which is not always well-conditioned. It is the purpose of this appendix to describe the new computational procedure which was used effectively to produce the data presented in Chapter 3.

Let m be the total number of measurements made at any particular check-point including the recording of the clock reading. Four such measurements are sufficient to determine the spaceship position and the clock correction. However, since the measurements are subject to error, additional redundant measurements can be employed to reduce the uncertainties in the fix. Let r be the number of these redundant measurements.

The sighting, which is to start at the reference time T , actually begins at the time $T - \delta T$ due to an error in the clock. It is shown in R-235 that for purposes of an error analysis no loss in generality results from the assumption that the sighting process

consumes no time. For each angular measurement made, a deviation δA from the reference value will result due to a displacement $\delta \underline{r}$ of the spaceship from its reference position at time T . If we define two column vectors $\delta \underline{A}_m$ and $\delta \underline{r}_4$ of dimensions m and 4 , respectively, by

$$\delta \underline{A}_m = \text{col}(\delta A_1, \delta A_2, \delta A_3, \delta T, \delta A_5, \dots, \delta A_m), \quad (\text{B-1})$$

$$\delta \underline{r}_4 = \text{col}(\delta \underline{r}, \delta T), \quad (\text{B-2})$$

then, as is shown in R-235, we have

$$\delta \underline{A}_m = U_{m4} \delta \underline{r}_4, \quad (\text{B-3})$$

where U_{m4} is a rectangular matrix with m rows and 4 columns whose elements depend only upon the geometry of the configuration and upon the velocity vectors of the spaceship and planets. It will be found convenient to partition this matrix as indicated by

$$U_{m4} = \begin{bmatrix} U_{44} \\ U_{r4} \end{bmatrix} \quad (\text{B-4})$$

where U_{44} is a four-dimensional square matrix and U_{r4} is a rectangular matrix with a number of rows equal to the number of redundant measurements. Note that the last row of U_{44} consists of all zero elements with the exception of the right-most element which is unity.

We must now distinguish between the measured or inferred values of the various quantities and their true values; e.g., $\delta \tilde{A}_m$ will be the measured value of the deviation from reference quantities and $\delta \underline{A}_m$ will be the true value of the deviation. The best direct measurement for $\delta \tilde{T}$ is, of course, zero, so that

$$\delta \tilde{A}_m = \text{col}(\delta \tilde{A}_1, \delta \tilde{A}_2, \delta \tilde{A}_3, 0, \delta \tilde{A}_4, \dots, \delta \tilde{A}_m), \quad (\text{B-5})$$

$$\delta \underline{\tilde{r}}_4 = \text{col} (\delta \underline{\tilde{r}}, \delta \underline{\tilde{T}}). \quad (\text{B-6})$$

One approach to the problem of determining an estimate for the vector $\delta \underline{\tilde{r}}_4$ is the method of maximum likelihood. If we assume that the measurement errors have a multidimensional Gaussian distribution with zero means and with a moment matrix $\underline{\Phi}_{mm}$, then the maximum likelihood estimate requires that $\delta \underline{\tilde{r}}_4$ be determined as the solution of the following set of four linear algebraic equations

$$U_{4m} \underline{\Phi}_{mm}^{-1} U_{m4} \delta \underline{\tilde{r}}_4 = U_{4m} \underline{\Phi}_{mm}^{-1} \delta \underline{\tilde{A}}_m, \quad (\text{B-7})$$

where U_{4m} is the matrix transpose of U_{m4} .

Once an estimate for $\delta \underline{\tilde{T}}$ has been obtained, a correction may be applied to $\delta \underline{\tilde{r}}$ to account for the fact that the time of starting the fix is in error. To obtain the deviations at the reference time T , indicated by $\delta \underline{\tilde{r}}_4(T)$, it is necessary to add to $\delta \underline{\tilde{r}}$ the vector distance travelled by the spaceship in time $\delta \underline{\tilde{T}}$. If we define a four-dimensional matrix X_{44} as

$$X_{44} = \begin{bmatrix} I_{33} & \underline{v}_s \\ 0 & 1 \end{bmatrix}, \quad (\text{B-8})$$

where I_{33} is the three-dimensional identity matrix and \underline{v}_s is the vector velocity of the spaceship at time T , then we have

$$\delta \underline{\tilde{r}}_4(T) = X_{44} \delta \underline{\tilde{r}}_4, \quad (\text{B-9})$$

and from Eq (B-7) we have

$$\delta \underline{\tilde{r}}_4(T) = X_{44} (U_{4m} \underline{\Phi}_{mm}^{-1} U_{m4})^{-1} U_{4m} \underline{\Phi}_{mm}^{-1} \delta \underline{\tilde{A}}_m. \quad (\text{B-10})$$

In our theory we have assumed that the individual measurement errors constitute a set of statistically independent random variables so that the moment matrix $\underline{\Phi}_{mm}$ is diagonal, with each

diagonal element being the variance of the corresponding measurement error. Therefore, we may partition this matrix as indicated by

$$\tilde{\Phi}_{mm} = \begin{bmatrix} \tilde{\Phi}_{44} & O_{4r} \\ O_{r4} & \tilde{\Phi}_{rr} \end{bmatrix}, \quad (B-11)$$

Now if we define the two square matrices*

$$P_{44} = U_{44}^{-1} \tilde{\Phi}_{44} U_{44}^{T-1}, \quad (B-12)$$

$$Q_{rr} = \tilde{\Phi}_{rr} + U_{r4} P_{44} U_{4r}, \quad (B-13)$$

it can be verified directly that

$$(U_{4m} \tilde{\Phi}_{mm}^{-1} U_{m4})^{-1} = P_{44} - P_{44} U_{4r} Q_{rr}^{-1} U_{r4} P_{44}. \quad (B-14)$$

Then by substituting (B-14) and (B-11) into Eq (B-10), we obtain, after a little manipulation,

$$\delta \tilde{r}_4(T) = X_{44} U_{44}^{-1} \left\{ \begin{bmatrix} I_{44} & O_{4r} \\ O_{r4} & I_{rr} \end{bmatrix} + B_{4r} Q_{rr}^{-1} \begin{bmatrix} -A_{r4} & I_{rr} \end{bmatrix} \right\} \delta \tilde{A}_m, \quad (B-15)$$

where we have defined the two rectangular matrices

$$A_{r4} = U_{r4} U_{44}^{-1}, \quad (B-16)$$

$$B_{4r} = \tilde{\Phi}_{44}^{-1} A_{4r}. \quad (B-17)$$

The matrix Q_{rr} may be expressed in terms of A_{r4} and B_{4r} by

$$Q_{rr} = \tilde{\Phi}_{rr} + A_{r4} B_{4r}. \quad (B-18)$$

* The superscript T on a matrix is used to denote the matrix transpose.

The relation between $\delta \tilde{r}_4(T)$ and $\delta \tilde{A}_m$ as expressed by Eq (B-15) has a significant advantage over the corresponding relationship given in Eq (B-10). In the former case no computational difficulty is encountered by assuming that any of the individual measurements is perfect. This assumption would lead to a singular \tilde{Z}_{mm} matrix which would be unmanageable in Eq (B-10) but would have no effect in the alternate expression. Furthermore, Eq (B-15) displays explicitly the effect of adding redundant measurements.

In the computation indicated in Eq (B-15) through (B-18), the inversion of two matrices U_{44} and Q_{rr} is required. Because of the special nature of U_{44} , i.e.,

$$U_{44} = \begin{pmatrix} u_{11} & u_{12} & u_{13} & u_{14} \\ u_{21} & u_{22} & u_{23} & u_{24} \\ u_{31} & u_{32} & u_{33} & u_{34} \\ 0 & 0 & 0 & 1 \end{pmatrix} = \left[\begin{array}{ccc|c} & & & u_{14} \\ & U_{33} & & u_{24} \\ & & & u_{34} \\ \hline 0 & 0 & 0 & 1 \end{array} \right], \quad (B-19)$$

the calculation of the inverse involves only a 3 x 3 matrix. We have

$$U_{44}^{-1} = \left[\begin{array}{ccc|c} & U_{33}^{-1} & -U_{33}^{-1} \begin{pmatrix} u_{14} \\ u_{24} \\ u_{34} \end{pmatrix} \\ \hline 0 & 0 & 0 & 1 \end{array} \right]. \quad (B-20)$$

The order of the matrix Q_{rr} is equal to the number of redundant measurements. Since it is composed of the sum of two matrices, one of which is diagonal, the inversion should proceed without difficulty except in rather specialized or artificial cases.

It remains to discuss the inversion of the U_{33} matrix defined in Eq (B-19). A particularly useful set of three observations consists of measuring the angles between a nearby planet and (1) a star, (2) a different star, and (3) the Sun. As we shall see, if

these three measurements are selected as the primary set and if the nearby planet is much closer than the Sun, the U_{33} matrix may not be inverted by straightforward means. In order to devise a procedure to handle this case, we must examine the elements of the U_{33} matrix.

For this purpose let unit row vectors \underline{n}_1 , \underline{n}_2 , \underline{n}_3 , and \underline{n}_4 be defined as follows:

- \underline{n}_1 , in the plane of the planet-star₁ measurement and normal to the direction from the spaceship to the planet;
- \underline{n}_2 , in the plane of the planet-star₂ measurement and normal to the direction from the spaceship to the planet;
- \underline{n}_3 , in the plane of the planet-Sun measurement and normal to the direction from the spaceship to the planet; and
- \underline{n}_4 , in the plane of the planet-Sun measurement and normal to the direction from the spaceship to the Sun.

Now if r and z are the distances to the Sun and the planet, respectively, then, according to the results of R-235, the U_{33} matrix is given by

$$U_{33} = \begin{pmatrix} \underline{n}_1/z \\ \underline{n}_2/z \\ \underline{n}_3/z + \underline{n}_4/r \end{pmatrix}. \quad (\text{B-21})$$

Since \underline{n}_1 , \underline{n}_2 , and \underline{n}_3 are each perpendicular to the same direction, they cannot be independent. Therefore, if $z \ll r$, the matrix U_{33} will have a small vanishing determinant.

To counter this difficulty, we define two scalar quantities a and b such that

$$\underline{n}_3 = a \underline{n}_1 + b \underline{n}_2, \quad (\text{B-22})$$

whereby Eq (B-21) may be written as

$$U_{33} = \begin{pmatrix} z^{-1} & 0 & 0 \\ 0 & z^{-1} & 0 \\ a z^{-1} & b z^{-1} & r^{-1} \end{pmatrix} \begin{pmatrix} \underline{n}_1 \\ \underline{n}_2 \\ \underline{n}_4 \end{pmatrix}. \quad (\text{B-23})$$

The coefficients a and b may be determined from

$$a = \frac{(\underline{n}_1 \cdot \underline{n}_3) - (\underline{n}_1 \cdot \underline{n}_2)(\underline{n}_2 \cdot \underline{n}_3)}{1 - (\underline{n}_1 \cdot \underline{n}_2)^2}, \quad (\text{B-24})$$

$$b = \frac{(\underline{n}_2 \cdot \underline{n}_3) - (\underline{n}_1 \cdot \underline{n}_2)(\underline{n}_1 \cdot \underline{n}_3)}{1 - (\underline{n}_1 \cdot \underline{n}_2)^2}. \quad (\text{B-25})$$

Since the vectors \underline{n}_1 , \underline{n}_2 , and \underline{n}_4 are independent, the inversion of the second matrix factor in Eq (B-23) presents no problem. The inversion of the first factor may be obtained explicitly. Therefore, we have

$$U_{33}^{-1} = \begin{pmatrix} \underline{n}_1 \\ \underline{n}_2 \\ \underline{n}_4 \end{pmatrix}^{-1} \begin{pmatrix} z & 0 & 0 \\ 0 & z & 0 \\ -a r & -b r & r \end{pmatrix}, \quad (\text{B-26})$$

as the required result.

Finally, in order to study the effects on navigation of the errors in the primary measurements, we write

$$\delta \tilde{\underline{A}}_m = \delta \underline{A}_m + \underline{\alpha}_m, \quad (\text{B-27})$$

$$\delta \tilde{\mathbf{r}}_4(T) = \delta \mathbf{r}_4(T) + \underline{\epsilon}_4, \quad (\text{B-28})$$

where α_m and ϵ_4 are the errors made in the measurements and the estimates. The components of the error vectors are denoted by

$$\underline{\alpha}_m = \text{col}(\alpha_1, \alpha_2, \alpha_3, -\delta T, \alpha_5, \dots, \alpha_m), \quad (\text{B-29})$$

$$\underline{\epsilon}_4 = \text{col}(\epsilon_1, \epsilon_2, \epsilon_3, \tau). \quad (\text{B-30})$$

By substituting Eq (B-27) and (B-28) into (B-15) and noting that the true deviations satisfy Eq (B-15) identically, we have

$$\underline{\epsilon}_4 = \mathbf{F}_{4m} \underline{\alpha}_m, \quad (\text{B-31})$$

where

$$\mathbf{F}_{4m} = \mathbf{X}_{44} \mathbf{U}_{44}^{-1} \left\{ \left\| \begin{array}{cc} \mathbf{I}_{44} & \mathbf{O}_{4r} \end{array} \right\| + \mathbf{B}_{4r} \mathbf{Q}_{rr}^{-1} \left\| \begin{array}{c} \mathbf{I}_{rr} \\ \mathbf{A}_{r4} \end{array} \right\| \right\}. \quad (\text{B-32})$$

Therefore, the correlation matrix \mathbf{E}_{44} of the estimate errors is obtained as

$$\mathbf{E}_{44} = \overline{\underline{\epsilon}_4 \underline{\epsilon}_4^T} = \mathbf{F}_{4m} \overline{\Phi}_{mm} \mathbf{F}_{m4}, \quad (\text{B-33})$$

where $\overline{\Phi}_{mm}$ is a diagonal matrix whose i^{th} element is $\overline{\alpha_i^2}$.

APPENDIX C

VARIABLE-TIME-OF-ARRIVAL
NAVIGATION THEORY

by

Richard H. Battin

TABLE OF CONTENTS

	Page
Appendix C	37
Variable-Time-of-Arrival Navigation Theory	
I. The Fundamental Navigation Equation	38
II. Analysis of the Velocity Correction	43
III. Navigation Error Analysis	48

APPENDIX C

VARIABLE-TIME-OF-ARRIVAL NAVIGATION THEORY

The navigation theory to be described in this appendix has as its primary objective that of minimizing the fuel consumption of the micro-rocket system without degrading the overall accuracy of the mission. The reduction in fuel requirement is accomplished by permitting the time of contact with the target planet to be a variable chosen in such a way that the velocity correction at any check-point will have the smallest possible magnitude. Just as in the fixed-time-of-arrival navigation scheme, the spaceship is controlled in the vicinity of a reference interplanetary trajectory. When the main propulsion stages of the booster rocket are completed, the vehicle proceeds in a solar orbit with an inaccuracy in the initial velocity attributable to injection guidance errors. At each of a number of check-points during the flight, positional deviations from the precomputed reference trajectory are determined from celestial observations. From these data velocity corrections are computed. By allowing a certain degree of flexibility in the exact time of arrival at the destination planet, only a fraction of the velocity correction needed to direct the vehicle toward the reference arrival point need be applied.

It is our objective here first to derive an expression for the appropriate velocity correction in terms of positional deviations from a reference path which is suitable for use by the spaceship. Following this, we shall determine explicitly expressions for the velocity corrections in terms of both the measurement and accelerometer errors. Then, for the purpose of an error analysis of

guidance accuracy, the final position and velocity errors will be related to the measurement and accelerometer errors.

I. The Fundamental Navigation Equation

Let $\underline{r}_s(t)$ and $\underline{v}_s(t)$ denote the position and velocity vectors of the spaceship in an inertial coordinate system with origin at the Sun, and let $\underline{g}(\underline{r}_s, t)$ denote the gravitational acceleration at position \underline{r}_s and time t . Then

$$\frac{d\underline{r}_s}{dt} = \underline{v}_s, \quad \frac{d\underline{v}_s}{dt} = \underline{g}(\underline{r}_s, t) \quad (C-1)$$

are the basic equations of motion of the spaceship except for those brief periods during which propulsion is applied.

Let the vectors $\underline{r}_o(t)$ and $\underline{v}_o(t)$ represent the position and velocity at time t associated with the prescribed reference trajectory, and define

$$\delta\underline{r}(t) = \underline{r}_s(t) - \underline{r}_o(t), \quad \delta\underline{v}(t) = \underline{v}_s(t) - \underline{v}_o(t). \quad (C-2)$$

Then, the deviations $\delta\underline{r}$ and $\delta\underline{v}$ may be approximately related by means of the linearized differential equations:

$$\frac{d(\delta\underline{r})}{dt} = \delta\underline{v}, \quad \frac{d(\delta\underline{v})}{dt} = \underline{R}_o(\underline{r}_o, t) \delta\underline{r}, \quad (C-3)$$

where $\underline{R}_o(\underline{r}_o, t)$ is a matrix whose elements are the partial derivatives of the components of $\underline{g}(\underline{r}_o, t)$ with respect to the components of \underline{r}_o .

A particularly useful fundamental set of solutions of Eq (C-3) may be developed in the following way. Let T_L and T_A be, respectively, the time of launch and the time of arrival at the destination planet. Then, define the matrices $\underline{R}(t)$, $\underline{R}^*(t)$, $\underline{V}(t)$, $\underline{V}^*(t)$ as the solutions of the matrix differential equations

$$\frac{dR}{dt} = V, \quad \frac{dR^*}{dt} = V^*, \quad (C-4)$$

$$\frac{dV}{dt} = R_O R, \quad \frac{dV^*}{dt} = R_O R^*,$$

which satisfy the initial conditions

$$\begin{aligned} R(T_L) &= O, & R^*(T_A) &= O, \\ V(T_L) &= I, & V^*(T_A) &= I. \end{aligned} \quad (C-5)$$

Here O and I denote, respectively, the zero and identity matrix. If we now write

$$\delta \underline{r}(t) = R(t) \underline{c} + R^*(t) \underline{c}^*, \quad (C-6)$$

$$\delta \underline{v}(t) = V(t) \underline{c} + V^*(t) \underline{c}^*, \quad (C-7)$$

where \underline{c} and \underline{c}^* are arbitrary constant vectors, it follows that these expressions satisfy the perturbation differential Eq (C-3) and contain precisely the required number of unspecified constants to meet any valid set of initial or boundary conditions.

Assume that measured positional deviations $\delta \tilde{\underline{r}}_{n-1}$ and $\delta \tilde{\underline{r}}_n$ from corresponding reference values are available at the times T_{n-1} and T_n of two successive fixes. Then Eq (C-6) may be written twice with T_{n-1} and T_n substituted for t . Solving this set for \underline{c} and \underline{c}^* and substituting these values into Eq (C-7), we have*

$$\delta \tilde{\underline{v}}_n = (B_n + B_n^*) \delta \tilde{\underline{r}}_n + (\Gamma_n + \Gamma_n^*) \delta \tilde{\underline{r}}_{n-1}, \quad (C-8)$$

where we have used the notation $R_n \equiv R(T_n)$, etc., and defined, for convenience, the matrices

*The superscripts - and + are used to distinguish the velocity just prior to a correction from the velocity immediately following the correction.

$$\begin{aligned}
A_n &= R_{n-1} R_n^{-1}, & C_n &= V_n R_n^{-1}, \\
\Gamma_n &= C_n (A_n - A_n^*)^{-1}, & B_n &= -\Gamma_n A_n^*,
\end{aligned} \tag{C-9}$$

with similar definitions for A_n^* , C_n^* , Γ_n^* , B_n^* obtained from Eq (C-9) by replacing all starred matrices by the corresponding unstarred ones and conversely.

Eq (C-8) provides a means of estimating the spaceship velocity at time T_n from positional information at times T_n and T_{n-1} . For fixed-time-of-arrival navigation there must be added to this velocity a calculated velocity increment $\tilde{\Delta}_n$ to arrive at the point $\underline{r}_O(T_A)$ at the time T_A . If the spaceship arrives at the reference point from its present position, there will be a velocity deviation $\delta\tilde{\underline{v}}(T_A)$ upon arrival which is related to $\delta\tilde{\underline{r}}_n$ by

$$\delta\tilde{\underline{r}}_n = R_n^* \delta\tilde{\underline{v}}(T_A).$$

The corresponding velocity deviation at time T_n ,

$$V_n^* \delta\tilde{\underline{v}}(T_A) = V_n^* R_n^{*-1} \delta\tilde{\underline{r}}_n = C_n^* \delta\tilde{\underline{r}}_n,$$

is precisely that which must be established at time T_N . Hence, the fixed-time-of-arrival required velocity correction is given by

$$\tilde{\Delta}_n = C_n^* \delta\tilde{\underline{r}}_n - \delta\tilde{\underline{v}}_n^- = H_n \delta\tilde{\underline{r}}_n - P_n \delta\tilde{\underline{r}}_{n-1}, \tag{C-10}$$

where the matrices H_n and P_n are defined by

$$H_n = C_n^* - (B_n + B_n^*), \quad P_n = \Gamma_n + \Gamma_n^*. \tag{C-11}$$

In order to calculate the variable-time-of-arrival required velocity correction, let us consider the effect of changing the arrival time T_A by a small amount δT . Let $\underline{r}_p(t)$ and $\underline{v}_p(t)$ be, respectively, the position and velocity vectors of the target planet.

Then the new point of contact will be $\underline{r}_p(T_A + \delta T)$, and associated therewith will be a somewhat different reference path. Let $\delta \underline{v}_o(T_L)$ be the vector change in launch velocity from the old reference trajectory which is needed to establish the spaceship in the new reference path. From the definition of the R matrix, it follows that at time T_A the spaceship position will be

$$\underline{r}_o(T_A) + R_A \delta \underline{v}_o(T_L).$$

At a time δT later the spaceship position will be

$$\underline{r}_o(T_A) + R_A \delta \underline{v}_o(T_L) + \underline{v}_o(T_A) \delta T,$$

and the corresponding planet position will be

$$\underline{r}_p(T_A) + \underline{v}_p(T_A) \delta T.$$

Assuming contact to be made at time $T_A + \delta T$, these positions are the same and we may solve for $\delta \underline{v}_o(T_L)$ to obtain

$$\delta \underline{v}_o(T_L) = - R_A^{-1} \underline{v}_R(T_A) \delta T, \quad (C-12)$$

where

$$\underline{v}_R(T_A) = \underline{v}_s(T_A) - \underline{v}_p(T_A) \quad (C-13)$$

is the velocity of the spaceship relative to the planet at the nominal arrival time T_A .

Now at the n^{th} check-point, the vector differences in position $\delta \underline{r}_o(T_n)$ and velocity $\delta \underline{v}_o(T_n)$ between the old and new reference trajectories are simply

$$\delta \underline{r}_o(T_n) = R_n \delta \underline{v}_o(T_L), \quad \delta \underline{v}_o(T_n) = V_n \delta \underline{v}_o(T_L).$$

Hence, from Eq (C-12)

$$\delta \underline{r}_o(T_n) = - R_n R_A^{-1} \underline{v}_R(T_A) \delta T, \quad (C-14)$$

$$\delta \underline{y}_o(T_n) = - \underline{V}_n \underline{R}_A^{-1} \underline{y}_R(T_A) \delta T. \quad (C-15)$$

Let the measured deviation in position from the old reference path be $\delta \underline{\tilde{r}}_n$, while $\partial \underline{\tilde{r}}_n$ is the corresponding deviation from the new reference path. Then

$$\partial \underline{\tilde{r}}_n = \delta \underline{\tilde{r}}_n - \delta \underline{r}_o(T_n). \quad (C-16)$$

With similar definitions for velocity deviations, we have

$$\partial \underline{\tilde{y}}_n = \delta \underline{\tilde{y}}_n - \partial \underline{y}_o(T_n). \quad (C-17)$$

By following the same arguments which led to Eq (C-10), we find that the velocity correction $\underline{\tilde{\Delta}}_n'(\delta T)$ to reach the new point of contact is

$$\underline{\tilde{\Delta}}_n'(\delta T) = \underline{C}_n^* \partial \underline{\tilde{r}}_n - \partial \underline{\tilde{y}}_n. \quad (C-18)$$

Using Eq (C-16), (C-17), (C-14), (C-15), and (C-10), we may write this in the form

$$\underline{\tilde{\Delta}}_n'(\delta T) = \underline{\tilde{\Delta}}_n - \underline{\nu}_n \delta T, \quad (C-19)$$

where, for convenience, we have defined the vector $\underline{\nu}_n$ by

$$\underline{\nu}_n = \underline{\Lambda}_n \underline{R}_A^{-1} \underline{y}_R(T_A) \quad (C-20)$$

and the matrix $\underline{\Lambda}_n$ by

$$\underline{\Lambda}_n = \underline{V}_n - \underline{C}_n^* \underline{R}_n. \quad (C-21)$$

With the object of picking δT so as to minimize the magnitude of $\underline{\tilde{\Delta}}_n'(\delta T)$, clearly the best choice is that which will render $\underline{\tilde{\Delta}}_n'(\delta T)$ normal to $\underline{\nu}_n$. Calling this value $\delta \tilde{T}_A$, we have, from Eq (C-19),

$$\delta \ddot{\mathbf{T}}_A = \frac{\dot{\hat{\Delta}}_n \cdot \underline{\nu}_n}{\underline{\nu}_n \cdot \underline{\nu}_n} \quad (\text{C-22})$$

As a consequence, the velocity correction $\dot{\hat{\Delta}}_n'$ of smallest magnitude which will accomplish the mission is simply related to $\dot{\hat{\Delta}}_n$ by

$$\dot{\hat{\Delta}}_n' = M_n \dot{\hat{\Delta}}_n \quad (\text{C-23})$$

The matrix M_n is defined by

$$M_n = I - \underline{\nu}_n \underline{\nu}_n^T / \underline{\nu}_n \cdot \underline{\nu}_n \quad (\text{C-24})$$

and is, obviously, singular.

II. Analysis of the Velocity Correction

In order to provide a basis for the selection of check-points, we will derive a relationship which will show explicitly how the velocity correction at time T_n is related to the initial velocity errors at launch, the errors associated with the positional fix, and the errors in establishing the desired velocity corrections at the previous check-points. For this purpose let $\underline{\Delta}_n'$ and $\underline{\eta}_n$ denote, respectively, the actual velocity applied at time T_n and the error made in the application of the desired correction $\dot{\hat{\Delta}}_n'$. Thus

$$\underline{\Delta}_n' = \underline{\Delta}_n + \underline{\eta}_n \quad (\text{C-25})$$

Similarly, we define $\underline{\epsilon}_n$ and $\underline{\delta}_n$, respectively, as the vector difference between the inferred and actual position and velocity deviations at time T_n , i.e.,

$$\delta \ddot{\mathbf{r}}_n = \delta \underline{\mathbf{r}}_n + \underline{\epsilon}_n, \quad \delta \ddot{\mathbf{v}}_n = \delta \underline{\mathbf{v}}_n + \underline{\delta}_n \quad (\text{C-26})$$

Therefore, it follows from Eq(C-25) and (C-10) that

$$\underline{\Delta}_n' + \underline{\eta}_n = M_n \left[C_n^* (\delta \underline{\mathbf{r}}_n + \underline{\epsilon}_n) - (\delta \underline{\mathbf{v}}_n + \underline{\delta}_n) \right] \quad (\text{C-27})$$

However, from Eq (C-8) and (C-26) we have

$$\underline{\delta}_n = (B_n + B_n^*) \underline{\epsilon}_n + (\underline{r}_n + \underline{r}_n^*) \underline{\epsilon}_{n-1}. \quad (C-28)$$

Thus, Eq (C-27) may be written as

$$\underline{\Delta}'_n = M_n (C_n^* \underline{\delta r}_n - \underline{\delta y}_n^-) + M_n (H_n \underline{\epsilon}_n - P_n \underline{\epsilon}_{n-1}) - \underline{\eta}_n. \quad (C-29)$$

It now remains to express $C_n^* \underline{\delta r}_n - \underline{\delta y}_n^-$ in terms of the error vectors at the present and previous check-points.

To this end we note that at time T_n^- we have

$$C_n^* \underline{\delta r}_n - \underline{\delta y}_n^- = - \underline{\Lambda}_n \underline{c},$$

which is obtained by premultiplying Eq (C-6) by C_n^* and subtracting Eq (C-7) with $t = T_n^-$. The constant vector \underline{c} is determined from Eq (C-6) and (C-7) with $t = T_{n-1}^*$. We find

$$\underline{c} = - \underline{\Lambda}_{n-1}^{-1} (C_{n-1}^* \underline{\delta r}_{n-1} - \underline{\delta y}_{n-1}^*). \quad (C-30)$$

Noting that

$$\underline{\delta y}_{n-1}^+ = \underline{\delta y}_{n-1}^- + \underline{\Delta}'_{n-1},$$

we have

$$C_n^* \underline{\delta r}_n - \underline{\delta y}_n^- = \underline{\Lambda}_n \underline{\Lambda}_{n-1}^{-1} (C_{n-1}^* \underline{\delta r}_{n-1} - \underline{\delta y}_{n-1}^- - \underline{\Delta}'_{n-1}). \quad (C-31)$$

Eq (C-31) may be used as a recursion formula so that by successive applications we have

$$C_n^* \underline{\delta r}_n - \underline{\delta y}_n^- = - \underline{\Lambda}_n \sum_{k=0}^{n-1} \underline{\Lambda}_k^{-1} \underline{\Delta}'_k, \quad (C-32)$$

if we define

$$\underline{\Delta}'_0 = \underline{\delta y}(T_L) = - \underline{\eta}_0$$

as the error in the initial launch velocity. Thus, Eq (C-29) becomes

$$\underline{\Delta}'_n = M_n(H_n \epsilon_n - P_n \epsilon_{n-1}) - \eta_n - M_n \wedge_n \sum_{k=0}^{n-1} \wedge_k^{-1} \underline{\Delta}'_k, \quad (C-33)$$

and $\underline{\Delta}'_n$ is expressed recursively in terms of all previous measurement and accelerometer errors.

In order to express $\underline{\Delta}'_n$ directly in terms of the errors, let us consider the following.

Lemma: If a sequence of vectors $\underline{a}_0, \underline{a}_1, \dots, \underline{a}_n$ is defined by

$$\begin{aligned} \underline{a}_0 &= \underline{b}_0 \\ \underline{a}_1 &= \underline{b}_1 + \psi_1 \Omega_0 \underline{a}_0 \\ &\vdots \\ \underline{a}_n &= \underline{b}_n + \psi_n \sum_{k=0}^{n-1} \Omega_k \underline{a}_k \quad n \geq 2 \end{aligned} \quad (C-34)$$

where $\underline{b}_0, \dots, \underline{b}_n; \psi_1, \dots, \psi_n$; and $\Omega_0, \dots, \Omega_{n-1}$ are arbitrary sequences of vectors and matrices, then

$$\underline{a}_n = \underline{b}_n + \psi_n \Omega_{n-1} \underline{b}_{n-1} + \psi_n \sum_{k=0}^{n-2} \prod_{j=n-1}^{k+1} (I + \Omega_j \psi_j) \Omega_k \underline{b}_k \quad n \geq 2 \quad (C-35)$$

Proof: If we define

$$\underline{d}_n = \sum_{k=0}^{n-1} \Omega_k \underline{a}_k \quad n \geq 2 \quad (C-36)$$

then we have

$$\underline{a}_n = \underline{b}_n + \psi_n \underline{d}_n$$

and it is sufficient to show that

$$\underline{d}_n = \Omega_{n-1} \underline{b}_{n-1} + \sum_{k=0}^{n-2} \prod_{j=n-1}^{k+1} (I + \Omega_j \psi_j) \Omega_k \underline{b}_k \quad n \geq 2. \quad (C-37)$$

For this purpose we note that, from the definitions of \underline{d}_n and \underline{a}_n , we have

$$\begin{aligned} \underline{d}_{n+1} &= \underline{d}_n + \Omega_{n-1} \underline{a}_{n-1} \\ &= \underline{d}_n + \Omega_{n-1} (\underline{b}_{n-1} + \psi_{n-1} \underline{d}_n) \\ &= (I + \Omega_{n-1} \psi_{n-1}) \underline{d}_n + \Omega_{n-1} \underline{b}_{n-1} \end{aligned}$$

as a difference equation for \underline{d}_n . Furthermore, it is a simple matter to verify that \underline{d}_n , as given by Eq (C-37), is the solution. For $n = 2$, Eq (C-37) yields

$$\underline{d}_2 = \Omega_1 \underline{b}_1 + (I + \Omega_1 \psi_1) \Omega_0 \underline{b}_0$$

which is clearly the same as that obtained from Eq (C-36), using the definitions of \underline{a}_0 and \underline{a}_1 . Hence, the lemma is proved.

=====

We may apply the lemma to the problem at hand by making the following identifications:

$$\begin{aligned} \underline{a}_n &\sim \underline{\Delta}'_n \\ \underline{b}_n &\sim M_n (H_n \underline{\epsilon}_n - P_n \underline{\epsilon}_{n-1}) \sim \underline{\eta}_n \\ \psi_n &\sim -M_n \Lambda_n \\ \Omega_k &\sim \Lambda_k^{-1}. \end{aligned}$$

Then if we define

$$X_{k,n} = \begin{cases} I & \text{for } k = n - 1 \\ \prod_{j=n-1}^{k+1} (I - \Delta_j^{-1} M_j \Delta_j) & \text{for } k \leq n - 2 \end{cases} \quad (C-38)$$

we may use Eq (C-35) to write Eq (C-33) in the form

$$\begin{aligned} \underline{\Delta}_n' &= M_n (H_n \underline{\epsilon}_n - P_n \underline{\epsilon}_{n-1}) - \eta_n \\ &- M_n \wedge_n \sum_{k=0}^{n-1} X_{k,n} \wedge_k^{-1} \left[M_k (H_k \underline{\epsilon}_k - P_k \underline{\epsilon}_{k-1}) - \underline{\eta}_k \right]. \end{aligned} \quad (C-39)$$

In the summation indicated in Eq (C-39), all terms for $k \leq n - 2$ have as a factor

$$M_n \wedge_n (I - \wedge_{n-1}^{-1} M_{n-1} \wedge_{n-1}).$$

Using Eq (C-20) and (C-24), this factor may be written as

$$\begin{aligned} &M_n \wedge_n \wedge_{n-1}^{-1} \underline{\nu}_{n-1} \underline{\nu}_{n-1}^T \wedge_{n-1} / \underline{\nu}_{n-1} \cdot \underline{\nu}_{n-1} \\ &= M_n \wedge_n \wedge_{n-1}^{-1} \wedge_{n-1} R_A^{-1} \underline{v}_R (T_A) \underline{v}_R^T (T_A) R_A^{T-1} \wedge_{n-1}^T \wedge_{n-1} / \underline{\nu}_{n-1} \cdot \underline{\nu}_{n-1} \\ &= M_n (\wedge_n R_A^{-1} \underline{v}_R (T_A) \underline{v}_R^T (T_A) R_A^{T-1} \wedge_n^T) \wedge_n^{T-1} \wedge_{n-1}^T \wedge_{n-1} / \underline{\nu}_{n-1} \cdot \underline{\nu}_{n-1} \\ &= M_n \underline{\nu}_n \underline{\nu}_n^T \wedge_n^{T-1} \wedge_{n-1}^T \wedge_{n-1} / \underline{\nu}_{n-1} \cdot \underline{\nu}_{n-1} \\ &= M_n (I - M_n) \wedge_n^{T-1} \wedge_{n-1}^T \wedge_{n-1} \left(\frac{\underline{\nu}_n \cdot \underline{\nu}_n}{\underline{\nu}_{n-1} \cdot \underline{\nu}_{n-1}} \right). \end{aligned}$$

Now, since

$$\left(\frac{\underline{\nu}_n \underline{\nu}_n^T}{\underline{\nu}_n \cdot \underline{\nu}_n} \right) \left(\frac{\underline{\nu}_n \underline{\nu}_n^T}{\underline{\nu}_n \cdot \underline{\nu}_n} \right) = \frac{\underline{\nu}_n \underline{\nu}_n^T (\underline{\nu}_n \cdot \underline{\nu}_n)}{(\underline{\nu}_n \cdot \underline{\nu}_n)^2} = \frac{\underline{\nu}_n \underline{\nu}_n^T}{\underline{\nu}_n \cdot \underline{\nu}_n}$$

it follows that

$$M_n (I - M_n) = O.$$

Hence, all terms in Eq (C-39) for $k \leq n - 2$ are identically zero, and we may write

$$\begin{aligned} \underline{\Delta}'_n &= M_n H_n \underline{\epsilon}_n - (M_n P_n + M_n \Lambda_n \Lambda_{n-1}^{-1} M_{n-1} H_{n-1}) \underline{\epsilon}_{n-1} \\ &\quad + M_n \Lambda_n \Lambda_{n-1}^{-1} M_{n-1} P_{n-1} \underline{\epsilon}_{n-2} - \underline{\eta}_n + M_n \Lambda_n \Lambda_{n-1}^{-1} \underline{\eta}_{n-1}. \end{aligned} \quad (C-40)$$

This relationship expresses explicitly the velocity impulse actually applied at time T_n in terms of present and previous measurement errors as well as errors in controlling the applied velocity. It is interesting to note that the corresponding expression for fixed-time-of-arrival navigation is obtained from Eq (C-40) by replacing the matrices M_n and M_{n-1} by the identity matrix.

III. Navigation Error Analysis

The effects of an initial velocity error, together with imperfect velocity corrections applied at the various check-points, will be: (1) a velocity deviation from the reference value at the destination planet; (2) a positional error or miss distance; and (3) a change from the scheduled arrival time. For use in the statistical analysis of guidance accuracy, we shall now develop appropriate expressions for each of these quantities in terms of the various measurement and accelerometer errors.

1) Velocity Deviation

We write Eq (C-6) and (C-7) for $t = T_n$ with $\delta \underline{r}(T_n) = 0$ and $\delta \underline{v}(T_n) = \underline{\Delta}'_n$, and solve for \underline{c} and \underline{c}^* to obtain

$$\underline{c} = \Lambda_n^{-1} \underline{\Delta}'_n, \quad \underline{c}^* = \Lambda_n^{*-1} \underline{\Delta}'_n.$$

Then, if $\delta \underline{v}_n(T_A)$ is the velocity deviation at time T_A due to the velocity impulse at time T_n , it follows from Eq (C-7) that

$$\delta \underline{v}_n(T_A) = V(T_A) \underline{c} + \underline{c}^*.$$

(Note that $V^*(T_A)$ is the identity matrix.) Now assuming N checkpoints, the total effect, obtained by superposition, is expressible as

$$\delta \underline{v}(T_A) = \sum_{n=0}^N \left[V(T_A) - R_n^{*-1} R_n \right] \Lambda_n^{-1} \underline{\Delta}_n', \quad (C-41)$$

where $\underline{\Delta}_0'$ is the initial velocity error at launch. By means of Eq (C-40), $\delta \underline{v}(T_A)$ may be expressed in terms of the errors $\underline{\epsilon}$ and $\underline{\eta}$.

2) Positional Error

In order to determine the positional deviation at the time of arrival, we use Eq (C-6) and (C-30) in the form

$$\delta \underline{r}(T_A) = R_A \Lambda_N^{-1} (\delta \underline{v}_N - C_N^* \delta \underline{r}_N + \underline{\Delta}_N'). \quad (C-42)$$

Then by substituting from Eq (C-32) with $n = N$, we obtain

$$\delta \underline{r}(T_A) = R_A \sum_{k=0}^N \Lambda_k^{-1} \underline{\Delta}_k'. \quad (C-43)$$

Only the component of $\delta \underline{r}(T_A)$ perpendicular to the direction of relative motion between the vehicle and the target planet is of interest in determining the actual miss distance. The other component along the direction of motion is more nearly responsible for an error in the scheduled arrival time. Denoting the actual miss distance vector by $\delta \underline{r}_a$, we have

$$\delta \underline{r}_a = M_a \delta \underline{r}(T_A), \quad (C-44)$$

where the matrix M_a is defined by

$$M_a = I - \underline{v}_R(T_A) \underline{v}_R^T(T_A) / \underline{v}_R(T_A) \cdot \underline{v}_R(T_A). \quad (C-45)$$

It would seem at first that the miss distance at the target planet is a function of measurement and velocity correction errors at all of the previous check-points. Indeed, as can be seen from Eq (C-43), the positional deviation from the reference arrival point does depend on all past errors; but only the measurement errors at the last two check-points, together with the last accelerometer error, affect the component $\delta \underline{r}_a$. For the proof we use Eq (C-33) with $n = N$ to write Eq (C-43) in the form

$$\begin{aligned} \delta \underline{r}(T_A) = & R_A \Lambda_N^{-1} M_N (H_N \underline{\epsilon}_N - P_N \underline{\epsilon}_{N-1}) - R_A \Lambda_N^{-1} \underline{\eta}_N \\ & + R_A (I - \Lambda_N^{-1} M_N \Lambda_N) \sum_{k=0}^{N-1} \Lambda_k^{-1} \underline{\Delta}_k. \end{aligned} \quad (C-46)$$

When we apply the transformation matrix M_a to this vector, the coefficient of the indicated summation can be shown to vanish identically. For if we use the definitions (C-20) and (C-24), we have

$$\begin{aligned} & M_a R_A (I - \Lambda_N^{-1} M_N \Lambda_N) \\ &= M_a R_A \Lambda_N^{-1} \underline{v}_N \underline{v}_N^T \Lambda_N / \underline{v}_N \cdot \underline{v}_N \\ &= M_a R_A \Lambda_N^{-1} \Lambda_N R_A^{-1} \underline{v}_R(T_A) \underline{v}_R^T(T_A) R_A^{T-1} \Lambda_N^{T-1} \Lambda_N / \underline{v}_N \cdot \underline{v}_N \\ &= M_a (I - M_a) R_A^{T-1} \Lambda_N^{T-1} \Lambda_N \left(\frac{\underline{v}_R(T_A) \cdot \underline{v}_R(T_A)}{\underline{v}_N \cdot \underline{v}_N} \right), \end{aligned}$$

and from the definition (C-45) of M_a it follows that

$$M_a (I - M_a) = O$$

Thus, the appropriate expression for the miss distance at the target planet is simply

$$\delta \underline{r}_a = M_a R_A \Lambda_N^{-1} \left[M_N (H_N \underline{\epsilon}_N - P_N \underline{\epsilon}_{N-1}) - \underline{\eta}_N \right]. \quad (C-47)$$

3) Change in the Scheduled Time of Arrival

From Eq (C-22) it is seen that at each check-point the optimum difference in arrival time from the nominal value T_A depends on the velocity correction $\underline{\Delta}_n$ which would be required to carry the vehicle to the nominal point of arrival $\underline{r}_p(T_A)$. If, at each of the previous check-points, the corrections have been of the variable time of arrival type, then $\underline{\Delta}_n$ will be a function of all previous measurement and velocity correction errors. The precise relationship is obtained as follows.

From Eq (C-10) and (C-26) we have

$$\underline{\Delta}_n = (C_n^* \delta \underline{r}_n - \delta \underline{v}_n) + H_n \underline{\epsilon}_n - P_n \underline{\epsilon}_{n-1},$$

and substituting from Eq (C-32) gives

$$\underline{\Delta}_n = H_n \underline{\epsilon}_n - P_n \underline{\epsilon}_{n-1} - \Lambda_n \sum_{k=0}^{n-1} \Lambda_k^{-1} \underline{\Delta}_k', \quad (C-48)$$

where $\underline{\Delta}_k'$ is expressed in terms of the $\underline{\epsilon}$'s and $\underline{\eta}$'s according to Eq (C-40). The final indicated change in the scheduled arrival time is obtained from Eq (C-22) with $n = N$. We find

$$\delta \tilde{T}_A = (\underline{v}_N \cdot \underline{v}_N)^{-1} \left\{ \underline{v}_N^T \left[H_N \underline{\epsilon}_N - P_N \underline{\epsilon}_{N-1} - \Lambda_N \sum_{k=0}^{N-1} \Lambda_k^{-1} \underline{\Delta}_k' \right] \right\}. \quad (C-49)$$

APPENDIX D

APPROXIMATE POSITION AND VELOCITY
OF THE MOON

by

Richard H. Battin

TABLE OF CONTENTS

	Page
Appendix D	55
Approximate Position and Velocity of the Moon	
Figure D-1, Geometry of the Moon's Position Relative	
to the Earth	59
Table D-1, Approximate Position of the Moon Compared	
with the Ephemeris Data for the Year 1958	60

APPENDIX D

APPROXIMATE POSITION AND VELOCITY OF THE MOON

In 1920 E. W. Brown published the most complete set of tables of the Moon's motion ever computed. The theory includes some 1500 separate terms which account for the perturbations in the motion of the Moon due to such effects as the presence of the Sun and planets, the ellipsoidal figure of the Earth, and the Moon's secular acceleration. Brown's tables have been used in the computations for the Nautical Almanac since the year 1923.

For the purposes of our navigational study, we have used a truncated form of Brown's series expansions to determine an approximate position of the Moon as a function of time. In this appendix we shall describe the computations used and compare the accuracy with ephemeris data. Following this, a procedure for obtaining the Moon's velocity vector will be outlined.

Let L, ϖ, Ω be the geocentric mean longitudes of the Moon, of its perigee and of its node. Let L' and ϖ' be the geocentric mean longitudes of the Sun and of its perigee. Further define

$$D = L - L', \quad \ell = L - \varpi, \quad \ell' = L' - \varpi', \quad \text{and} \quad F = L - \Omega \quad (\text{D-1})$$

to which have been associated the following names:

$D,$	half argument of the "variation",
$\ell,$	argument of the principal elliptic term,
$\ell',$	argument of the "annual equation"; and
$F,$	argument of the principal term in latitude.

Now if r is the distance of the Moon to the center of the Earth, v the true longitude measured in the plane of the ecliptic, and ϕ the latitude above the plane of the ecliptic, then it is known that r , $v - L$, and ϕ are expressible by sums of periodic terms whose arguments are algebraic sums of multiples of the four angles D , \mathcal{L} , \mathcal{L}' , and F . In place of the distance r it is customary to use a quantity called the "horizontal parallax" defined as half the angular diameter of the Earth as seen from the Moon.

In our particular approximation the expressions for the longitude, latitude, and parallax are taken from the Memoirs of the Royal Astronomical Society, Vol. LVII, pp. 109-145, and include only the effects of the solar perturbations. Furthermore, for longitude and latitude, only terms whose coefficients exceed 60 seconds of arc in magnitude were retained. For the parallax approximation, the lower limit was set at one second. The relevant expressions in seconds of arc are as follows:

$$\begin{aligned}
 \text{longitude} = & L + 22639.500 \sin \mathcal{L} \\
 & - 4586.426 \sin (\mathcal{L} - 2D) + 2369.902 \sin 2D \\
 & + 769.016 \sin 2\mathcal{L} - 668.111 \sin \mathcal{L}' \\
 & - 411.608 \sin 2F - 211.656 \sin (2\mathcal{L} - 2D) \\
 & - 205.962 \sin (\mathcal{L} + \mathcal{L}' - 2D) - 125.154 \sin D \\
 & + 191.953 \sin (\mathcal{L} + 2D) - 165.145 \sin (\mathcal{L}' - 2D) \\
 & + 147.693 \sin (\mathcal{L} - \mathcal{L}') - 109.667 \sin (\mathcal{L} + \mathcal{L}'); \\
 & \hspace{15em} (D-2)
 \end{aligned}$$

$$\begin{aligned}
 \text{latitude} = & 18461.480 \sin F + 1010.180 \sin (\mathcal{L} + F) \\
 & - 999.695 \sin (F - \mathcal{L}) - 623.658 \sin (F - 2D) \\
 & + 117.262 \sin (F + 2D) + 199.485 \sin (F + 2D - \mathcal{L}) \\
 & - 166.577 \sin (\mathcal{L} + F - 2D) + 61.913 \sin (2\mathcal{L} + F); \text{ and} \\
 & \hspace{15em} (D-3)
 \end{aligned}$$

$$\begin{aligned}
\text{parallax} = & 3422.7000 + 186.5398 \cos \ell \\
& + 34.3117 \cos (\ell - 2D) + 28.2333 \cos 2D \\
& + 10.1657 \cos 2\ell + 3.0861 \cos (\ell + 2D) \\
& + 1.9202 \cos (\ell' - 2D) + 1.4455 \cos (\ell + \ell' - 2D) \\
& + 1.1542 \cos (\ell - \ell'). \quad (D-4)
\end{aligned}$$

The fundamental arguments in these equations are functions of time and are given in Brown's Tables of the Motion of the Moon. The time t_c is expressed in Julian centuries of 36525 days, the epoch being 1900.0. Brown's equations for these quantities are:

$$\begin{aligned}
L = & 270^\circ 26' 11.71'' + 1336^r 307^\circ 53' 26.06'' t_c \\
& + 7.14'' t_c^2 + 0.0068'' t_c^3; \quad (D-5)
\end{aligned}$$

$$\begin{aligned}
\mathfrak{A} = & 334^\circ 19' 46.40'' + 11^r 109^\circ 02' 02.52'' t_c \\
& - 37.17'' t_c^2 = 0.045'' t_c^3; \quad (D-6)
\end{aligned}$$

$$\begin{aligned}
\Omega = & 259^\circ 10' 59.79'' - 5^r 134^\circ 08' 31.23'' t_c \\
& + 7.48'' t_c^2 + 0.008'' t_c^3; \quad (D-7)
\end{aligned}$$

$$\begin{aligned}
D = & 350^\circ 44' 23.67'' + 1236^r 307^\circ 07' 17.93'' t_c \\
& + 6.05'' t_c^2 + 0.0068'' t_c^3; \quad (D-8)
\end{aligned}$$

$$\begin{aligned}
\ell = & 296^\circ 06' 25.31'' + 1325^r 198^\circ 51' 23.54'' t_c \\
& + 44.31'' t_c^2 + 0.0518'' t_c^3; \quad (D-9)
\end{aligned}$$

$$\begin{aligned}
\ell' = & 358^\circ 28' 33.00'' + 99^r 359^\circ 02' 59.10'' t_c \\
& - 0.54'' t_c^2 - 0.0120'' t_c^3; \text{ and} \quad (D-10)
\end{aligned}$$

$$F = 11^{\circ} 15' 11.92'' + 1342^{\circ} 82' 1' 57.29'' t_c - 0.34'' t_c^2 - 0.0012'' t_c^3. \quad (D-11)$$

At Greenwich mean noon 1900.0 there have elapsed 2415020 Julian days so that t_c may be expressed in the form

$$t_c = \frac{\text{Julian Day Number} - 2415020}{36525} \quad (D-12)$$

(Note: The Julian day begins at noon while the Calendar day begins at midnight. For example, the calendar date January 7.0, 1958 is equivalent to the Julian date 2436210.5.)

In order to determine the accuracy of the approximate calculation, the longitude, latitude, and parallax were computed at 10-day intervals during the year 1958, beginning with January 7.0, and compared with data taken from The American Ephemeris and Nautical Almanac. The results are tabulated in Table D-1. The maximum deviation in longitude is just over 3 minutes of arc, in latitude, just over 2 minutes, and in parallax, just over 2 seconds of arc.

The velocity vector of the Moon is calculated by assuming that at each instant the Moon is moving in an elliptical path about the Earth with orbital elements obtained as follows: (1) mean longitude of the node Ω obtained from Eq. (D-7); (2) mean longitude of perigee ϖ obtained from Eq. (D-6); (3) mean distance $a = 238,840$ miles or 0.00256817 astronomical units; (4) mean eccentricity $e = 0.05473074$. These elements, together with the instantaneous lunar position, are sufficient for the calculation.

Referring to Fig. D-1, let x, y, z be a geocentric rectangular ecliptic-oriented coordinate system with the x axis in the direction of the vernal equinox. Then, if $\bar{i}, \bar{j}, \bar{k}$ are unit vectors in the direction of the coordinate axes, the unit vector \bar{u}_M in the direction of the Moon is obtained from the longitude and latitude according to

$$\bar{u}_M = \cos \phi \cos v \bar{i} + \cos \phi \sin v \bar{j} + \sin \phi \bar{k}. \quad (D-13)$$

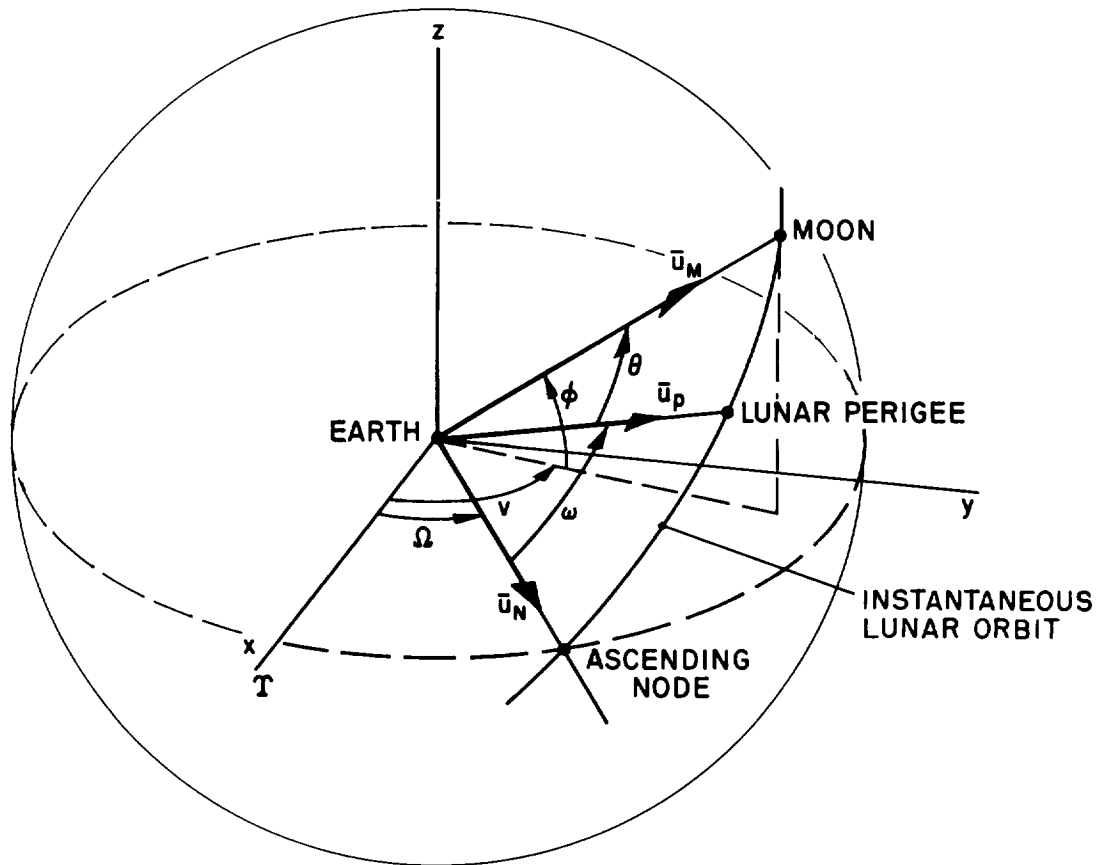


Fig. D-1 Geometry of moon's position relative to the Earth.

		APPROXIMATE POSITION						EPHEMERIS POSITION												
DATE		LONGITUDE			LATITUDE			HORIZONTAL PARALLAX			LONGITUDE			LATITUDE			HORIZONTAL PARALLAX			
		DEG	MN	SEC	DEG	MN	SEC	DEG	MN	SEC	DEG	MN	SEC	DEG	MN	SEC	DEG	MN	SEC	
JAN.	7	121	35	42.2	-5	00	14.0	59	29.50				121	35	16.4	-5	00	20.5	59	29.21
JAN.	17	261	40	54.5	+3	30	47.6	56	59.50				261	42	03.9	+3	31	13.6	56	57.65
JAN.	27	24	34	41.8	+0	59	23.2	54	25.17				24	34	29.5	+0	58	22.8	54	25.69
FEB.	6	159	51	18.7	-4	07	13.9	60	46.23				159	53	13.1	-4	06	13.2	60	46.58
FEB.	16	296	55	41.3	+4	59	30.6	55	21.82				296	57	49.1	+5	00	10.1	55	21.25
FEB.	26	56	57	00.1	-2	04	46.2	55	16.48				56	59	34.4	-2	06	18.7	55	17.06
MAR.	8	198	28	33.5	-1	14	00.7	60	55.20				198	30	59.7	-1	12	39.7	60	55.81
MAR.	18	330	38	45.8	+4	32	31.2	54	17.32				330	39	24.5	+4	32	27.0	54	16.89
MAR.	28	91	05	21.3	-4	31	27.2	56	39.22				91	08	31.5	-4	32	27.1	56	38.34
APR.	7	236	19	39.2	+2	14	09.4	59	54.00				236	19	57.1	+2	14	31.2	59	54.46
APR.	17	3	15	58.6	+2	29	55.6	53	58.59				3	15	13.6	+2	29	29.9	53	57.74
APR.	27	127	29	07.1	-5	15	40.7	58	10.56				127	30	22.8	-5	15	55.6	58	08.96
MAY	7	272	40	36.1	+4	38	09.1	58	09.54				272	40	07.0	+4	38	20.3	58	11.07
MAY	17	35	34	27.6	-0	23	41.6	54	27.93				35	34	07.2	-0	23	49.5	54	27.54
MAY	27	165	46	46.1	-3	41	27.7	59	15.58				165	47	22.7	-3	41	33.8	59	15.91
JUNE	6	307	16	49.7	+5	06	14.4	56	20.77				307	16	42.6	+5	07	13.1	56	23.25
JUNE	16	68	34	26.3	-3	10	40.5	55	42.77				68	35	29.0	-3	11	13.2	55	43.16
JUNE	26	204	57	58.6	-0	22	22.8	59	24.66				204	57	52.7	-0	22	23.1	59	26.63
JULY	6	340	17	50.1	+3	43	19.2	54	58.10				340	17	41.4	+3	44	21.8	54	59.60
JULY	16	103	09	44.8	-4	50	39.3	57	32.08				103	10	53.4	-4	51	47.6	57	32.45

TABLE D-1

COMPARISON OF APPROXIMATE POSITION OF THE MOON WITH EPHEMERIS DATA
FOR THE YEAR 1958

DATE	APPROXIMATE POSITION						EPHEMERIS POSITION									
	LONGITUDE			LATITUDE			LONGITUDE			LATITUDE						
	DEG	MN	SEC	DEG	MN	SEC	DEG	MN	SEC	DEG	MN	SEC				
JULY 26	243	47	48.6	+3	03	43.0	58	32.29	243	47	51.8	+3	03	51.4	58	33.40
AUG 5	12	15	27.2	+1	06	37.7	54	15.29	12	15	52.3	+1	07	13.8	54	15.40
AUG 15	139	38	11.0	-4	34	34.5	59	29.37	139	40	35.2	-4	34	52.5	59	29.56
AUG 25	281	16	07.7	+4	59	41.1	57	01.17	281	16	30.1	+4	59	42.5	57	01.40
SEPT 4	44	01	58.5	-1	52	38.7	54	15.06	44	03	26.1	-1	52	28.9	54	15.60
SEPT 14	177	36	36.9	-2	13	34.9	60	55.86	177	39	45.1	-2	12	00.8	60	55.82
SEPT 24	316	49	08.4	+4	47	18.6	55	27.80	316	49	23.4	+4	47	04.4	55	28.54
OCT 4	76	37	41.2	-4	18	01.1	54	56.16	76	38	58.3	-4	18	23.5	54	57.82
OCT 14	216	10	05.3	+1	14	24.4	61	17.78	216	12	11.9	+1	16	44.7	61	16.14
OCT 24	350	29	46.3	+2	48	29.0	54	22.70	350	29	10.3	+2	48	04.4	54	23.78
NOV 3	110	50	46.2	-5	15	35.4	56	11.36	110	51	04.6	-5	16	08.1	56	12.46
NOV 13	254	15	59.0	+4	06	40.6	60	26.34	254	16	10.3	+4	08	16.2	60	23.10
NOV 23	22	52	39.1	-0	05	07.7	54	01.05	22	52	43.8	-0	05	19.2	54	02.05
DEC 3	147	01	52.4	-4	09	15.4	57	40.37	147	02	05.6	-4	09	36.0	57	40.18
DEC 13	290	58	57.6	+5	03	44.5	58	43.25	290	58	50.2	+5	04	44.4	58	41.23
DEC 23	54	53	57.7	-2	54	26.7	54	26.37	54	54	38.8	-2	54	40.5	54	27.79

TABLE D-1

Further, we introduce two unit vectors \bar{u}_N and \bar{u}_P in the respective directions of the ascending node and the lunar perigee. If i is the angle of inclination of the lunar orbit with respect to the ecliptic and ω , the longitude of perigee measured from the ascending node, we find

$$\bar{u}_N = \cos \Omega \bar{i} + \sin \Omega \bar{j}, \quad (D-14)$$

$$\begin{aligned} \bar{u}_P = & (\cos \Omega \cos \omega - \sin \Omega \sin \omega \cos i) \bar{i} \\ & + (\sin \Omega \cos \omega + \cos \Omega \sin \omega \cos i) \bar{j} \\ & + \sin \omega \sin i \bar{k}, \end{aligned} \quad (D-15)$$

where

$$\begin{aligned} \cos i = & \frac{|\bar{u}_N \times \bar{u}_M \cdot \bar{k}|}{|\bar{u}_N \times \bar{u}_M|} \\ = & \frac{\cos \phi \sin |v - \Omega|}{\sqrt{1 - \cos^2 \phi \cos^2 (v - \Omega)}}. \end{aligned} \quad (D-16)$$

The true anomaly θ of the Moon may be obtained from Eq. (D-13) and (D-15) in the form

$$\begin{aligned} \cos \theta = & \bar{u}_P \cdot \bar{u}_M \\ = & \cos \phi \cos \omega \cos (v - \Omega) \\ & + \cos \phi \sin \omega \cos i \sin (v - \Omega) + \sin \phi \sin \omega \sin i. \end{aligned} \quad (D-17)$$

The sign of $\sin \theta$ is the same as the sign of

$$\bar{u}_P \times \bar{u}_M \cdot \bar{k}$$

so that

$$\sin \theta = \text{sgn} [\cos \omega \sin (v - \Omega) - \sin \omega \cos i \cos (v - \Omega)] \sqrt{1 - \cos^2 \theta}. \quad (D-18)$$

Now if a_E is the Earth's equatorial radius, taken to be 3963.35 miles, the distance r of the Moon from the center of the Earth is computed from

$$r = a_E / \sin (\text{Parallax}) . \quad (\text{D-19})$$

Finally, the polar components v_r , v_θ of the Moon's velocity vector in the plane of the lunar orbit are determined from

$$v_r = \text{sgn} (\sin \theta) \frac{na^2}{r} \sqrt{e^2 - \left(\frac{r}{a} - 1\right)^2}, \quad (\text{D-20})$$

$$v_\theta = \frac{na^2}{r} \sqrt{1 - e^2}, \quad (\text{D-21})$$

where n is the mean angular motion of the Moon about the Earth and may be computed from

$$\frac{2\pi}{n} = 27.321661 \text{ days}. \quad (\text{D-22})$$

Alternately, we have

$$na = 0.21572212 \text{ astronomical units/year}. \quad (\text{D-23})$$

The Moon's velocity vector in the reference ecliptic coordinate system may now be readily determined.

APPENDIX E

R-276

DESIGN PRINCIPLES FOR A GENERAL
CONTROL COMPUTER

by

R. Alonso

J. H. Laning, Jr.

March 1960

Note: This appendix is separately published as Instrumentation
Laboratory Report R-276, April, 1960.

ACKNOWLEDGMENT

Part of the work presented in this report was done by Dr. Richard H. Battin and Mr. Robert E. Oleksiak. Dr. Battin took part in the discussions on order codes and authored the final refined version of the "representative" order code. Mr. Oleksiak had a major part in the electrical design of both the fixed and erasable memory systems constructed and operated to date, designed the specialized and necessary Z and P registers, and contributed to the invention of the priority circuit.

This report was prepared under Project DSR 52-156 sponsored by the Ballistic Missile Division of the Air Research and Development Command through USAF Contract AF 04(647)-303 and Project DSR 53-138, Division of Sponsored Research, Massachusetts Institute of Technology, sponsored by the Bureau of Ordnance, Department of the Navy, under Contract NOrd 17366.

The cost of reproducing this report here was borne by DSR Project 55-171 sponsored by the National Aeronautics and Space Administration through Contract NASw-130.

ABSTRACT

A set of techniques is described which permits the design of computers particularly well suited for outer space and other airborne control applications. These techniques are unified by reference to a representative computer.

Some of the properties of the computer are: variable speed, power consumption proportional to speed, relatively few transistors, relatively large storage for program and constants, and parallel transfer of words. Certain features of the input system permit automatic incrementing of counters and automatic interruption of normal computer processes upon receipt of inputs.

The program and constants are stored in a wired-in form of memory which permits unusually high bit densities.

TABLE OF CONTENTS

	Page
Introduction	71
Chapter I General Organization of Storage	75
A. The Core Rope (Wired-in Memory)	75
B. Erasable Storage Registers	81
C. The Erasable Storage System	87
D. The Complete Memory System	90
Chapter II Execution of Computer Instructions	97
A. Central Registers and Overall Organization	97
B. The Sequence Generator	103
C. Representative Order Code System	107
D. Addition and Multiplication	124
Chapter III Input-Output Systems	131
A. Introduction	131
B. The Priority Circuit	132
C. The Automatic Interrupt	139
D. State Inputs	140
E. Counter Inputs	144
F. Output Devices	161
Chapter IV General Remarks	161
A. Sizes and Weights	161
B. Reliability Problems	162
C. Practical Experience to Date	167
D. Power and Speed	168

LIST OF ILLUSTRATIONS

Figure		Page
1-1	Word core rope	76
1-2	Selection of a core in a core rope	78
1-3	Timing of inhibiting, set and clear currents	79
1-4	Sensing wires on a core rope	80
1-5	Symbolic erasable register	82
1-6	Erasable register	84
1-7	Erasable storage system	86
1-8	System of ten erasable storage registers	88
1-9	Erasable storage selection system	91
1-10	Complete memory system	92
2-1	Over-all organization of the representative computer	98
2-2	Sequence generator	102
2-3	Control pulse generation by sequence generator	104
2-4	Representative central register system	108
2-5	Average number of cycles for addition	127
3-1	Priority circuit	133
3-2	Priority circuit system	135
3-3	Input circuits for priority circuit system	136
3-4	Unit M for input to priority circuit	141
3-5	Circuitry for state input	143
3-6	Ideal counter register	146
3-7	System for counter inputs	147
3-8	Counter inputs -- system for dealing with positive and negative increments.	151
3-9	Output switch circuit	153
3-10	Gating of a pulse generator by an output switch	154

Figure		Page
3-11	Generator of several pulse rates. Pulse rate $Q = a_1 R + a_2 \frac{R}{2} + a_3 \frac{R}{4} + a_4 \frac{R}{8} + a_5 \frac{R}{16}$, where $a_i = 0$ or 1 , and R is some pulse rate determined by the clock. The contents of an output register determine which a are 0 and which are 1	156
3-12	Details of an output rate register	157
3-13	Pulse rate generation and examples	159
4-1	Sixteen word erasable storage system. (Clearing tran- sistors at center. Write transistor-core combinations at edges. Write lines are under registers.)	163
4-2	Physical characteristics of components	165

DESIGN PRINCIPLES FOR A GENERAL CONTROL COMPUTER

INTRODUCTION

The subject of this report is a medium-speed, general - purpose digital computer designed to serve as a controlling element for a deep space probe or other complex airborne system; it is a continuation of the work presented in Report R-235^{*}. Major features of the design are the small size and weight of this computer, its rugged construction, low power consumption, and a combination of circuits and logical organization which allows many sophisticated operations to be accomplished quite simply. All of this can be achieved by circuits composed of magnetic cores, diodes, and transistors, with the latter being few in number relative to comparable existing computers.

The computer is not presented in the context of a specific application, but rather as a collection of useful techniques and approaches, especially with regard to inputs and outputs. To assist the reader in visualizing reasonable areas of potential uses, sizes and speeds for a representative machine are nonetheless discussed. These would vary with the control application, but are meaningful in indicating the kinds of problems the computer can best solve. The remainder of this Introduction summarizes the principal characteristics of this machine.

^{*}Chapter 10, Vol III, A Recoverable Interplanetary Space Probe
MIT Instrumentation Laboratory Report R-235, July 1959.

Computer memory consists of a wired-in section for program and constants plus a smaller volatile memory for variable items. Both are random access in nature, with all transfers occurring in parallel. The wired-in memory provides for storage of information at a density between ten and one hundred times that achieved with the volatile type of memory (see Chap. 4, Section A of this report). As a result, the proposed computer is well suited for the solution of a large class of airborne and spaceborne control problems which require the versatility afforded by long and intricate logical or arithmetic programs, or require the storage of many constants, but which use relatively few variables. The representative machine possesses about 4000 words of wired-in storage and 128 words of volatile storage, all words being of the order of 23 bits in length, plus a parity bit. Estimates of a representative computer's size and weight are 0.37 ft³ and 20 lb.

Various input-handling techniques are presented which assist efficient computer utilization. The input lines are not scanned; when an input line becomes active, the computer interrupts its present activity to tend to that input, returning to its former task after the input has been dealt with. This mode of operation is particularly advantageous in the case in which there are many slowly varying inputs. (A scanning system requires that computer speed be increased with an increase in the number of input lines.) However, the computer is also well equipped for the counting of rapid trains of input pulses; each such pulse costs only two clock cycles* on the part of the computer for its processing. These techniques employ the efficient time sharing of a small amount of common equipment and thereby permit a relatively complex control situation to be handled by a small computer.

* A clock cycle is composed of two clock pulses, α and β .

The order code for the representative computer is of the single address type and is one in which 3 to 5 clock cycles are required for the operations involving transfer of information from one point to another or for logical operations. Addition and multiplication, however, require an average of about 27 and 750 clock cycles respectively; in fact, multiplication is accomplished by a subroutine. At a nominal 100 KC clock rate, multiplication requires 7.5 ms, addition 0.27 ms, and many other operations only 0.04 ms.

Computer speed is variable, and can be controlled both by external events and by the computer program itself. When functioning at full speed, the computer will consume power on the order of 20 watts, exclusive of input and output equipment. The electrical design, however, permits the computer to function at any clock cycle rate from a nominal maximum of 100 KC to zero cps, with a power consumption proportional to speed. Hence it is possible to take advantage of long periods during which the computer need not operate very rapidly, and save power. While in this semi-dormant status, the ability of the computer to process input information instantly on demand is in no way impaired; thus the effective operation speed can be exactly that dictated by the input requirement. This property is clearly useful in the case of a deep space probe.

The computer is expected to be long-lived, in the sense of no catastrophic failures. This is because the number of active components is low (of the order of 1500 transistors), and the power dissipated in each of these is well within the components' ratings. The possible sophistication and length of the computer program can be used for checking computations and to provide alternative actions in the case of failures of parts of the controlled system. Thus the survival capabilities of the controlled system could be increased significantly, especially in the case of temporary malfunctions.

The remaining chapters in this report discuss in some detail the logic and circuits for this computer. An effort is made to lead the reader progressively from the simple and separate elements of which the machine is composed, toward the complexities of the overall logical organization; the report is therefore best read in sequence. Many details are passed over in the process; the first rudimentary version of the machine is just now being assembled, and many changes are certain to be made. The overall organization and logic, however, appear to be sufficiently unique and promising in their possible applications to justify a report at this time.

CHAPTER 1

GENERAL ORGANIZATION OF STORAGE

A. The Core Rope (Wired-In Memory)

The term "core rope" is used to denote the device used as the wired-in or "fixed" storage for this computer and as the selection switch controlling operation of the volatile or "erasable" storage. The term is derived from the physical appearance of the array of cores and wiring; a rope of 256 cores is shown in Fig. 1-1. By means of the rope wiring pattern, an arbitrary binary address can be translated into the switching of a single corresponding rope core, which may in turn cause other actions such as read-out of an associated erasable (volatile) storage register. Additional sensing-wiring through the core permits the read-out of a wired-in binary number upon resetting the core, providing a medium for fixed storage.

The core rope, which was invented by Olsen of Lincoln Laboratory, and which is also known as the Diamond Switch, the Rajchman Selection Switch, and the Linear Selection Switch, makes use of cores as saturable pulse transformers. The cores are small permalloy ribbon cores, with a fairly square hysteresis loop. A small 8-core rope is shown in schematic form in Fig. 1-2. There are two pulse generators, one providing a ground at point (1) at TIME 1, and the other providing a ground at point (2) at TIME 2. The duration and timing of these grounds are shown in Fig. 1-3. The lines marked A, \bar{A} , B, \bar{B} , C, and \bar{C} are called Inhibitors. The slanted line at the intersection of a core and a wire shows that the wire in question threads that core. The

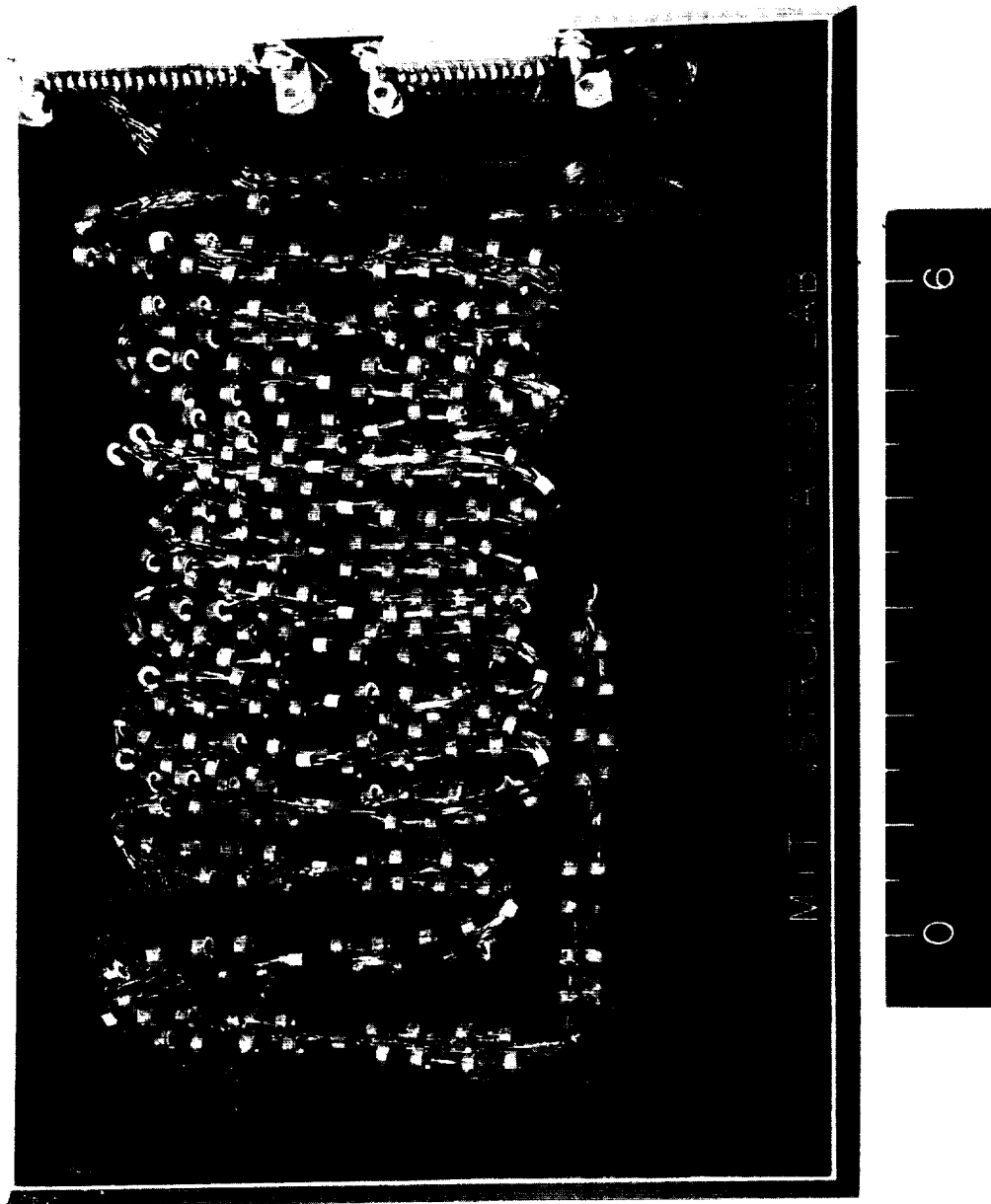


Fig. 1-1 Word core rope.

direction of the slant relative to the current in the wire indicates the direction of threading.

All eight cores shown in Fig. 1-2 are initially in state ZERO. The Set line puts a current through all the cores in a sense that drives all cores toward state ONE. A current-carrying Inhibitor, on the other hand, prevents those cores which it threads from switching from ZERO to ONE. If it does not carry current, i. e. , if it is not connected to point (1) by virtue of the position of the appropriate switch^{*}, then it has no effect on any core. Each pair of Inhibitors in Fig. 1-2 may be regarded as one binary digit position in a 3-digit address; one and only one Inhibitor from a pair is active at any time. It is easy to verify from the wiring pattern shown that each of the eight possible combinations of Inhibitors connected to the pulse source corresponds to a single core being switched from ZERO to ONE by a current through the Set line, all other cores being inhibited by at least one inhibiting current.

The Clear line, which carries current at TIME 2, tends to drive all the cores toward state ZERO. However, there is only one core to switch, the core that is in state ONE at that time.

Thus far we have considered how a core is selected, but not how information is stored in that core. This is accomplished by means of an additional set of windings on the cores of the rope, as shown in Fig. 1-4. As before, a threaded core is shown by means of a slanted line at the intersection of core and wire. For example, core number (7) is threaded only by Sense line d; core (6) is threaded by all Sense lines. Upon selecting and switching a core (say (4)), all Sense lines which thread it will

^{*}Symbolized in Fig. 1-2 as a relay contact, but electronic in practice.

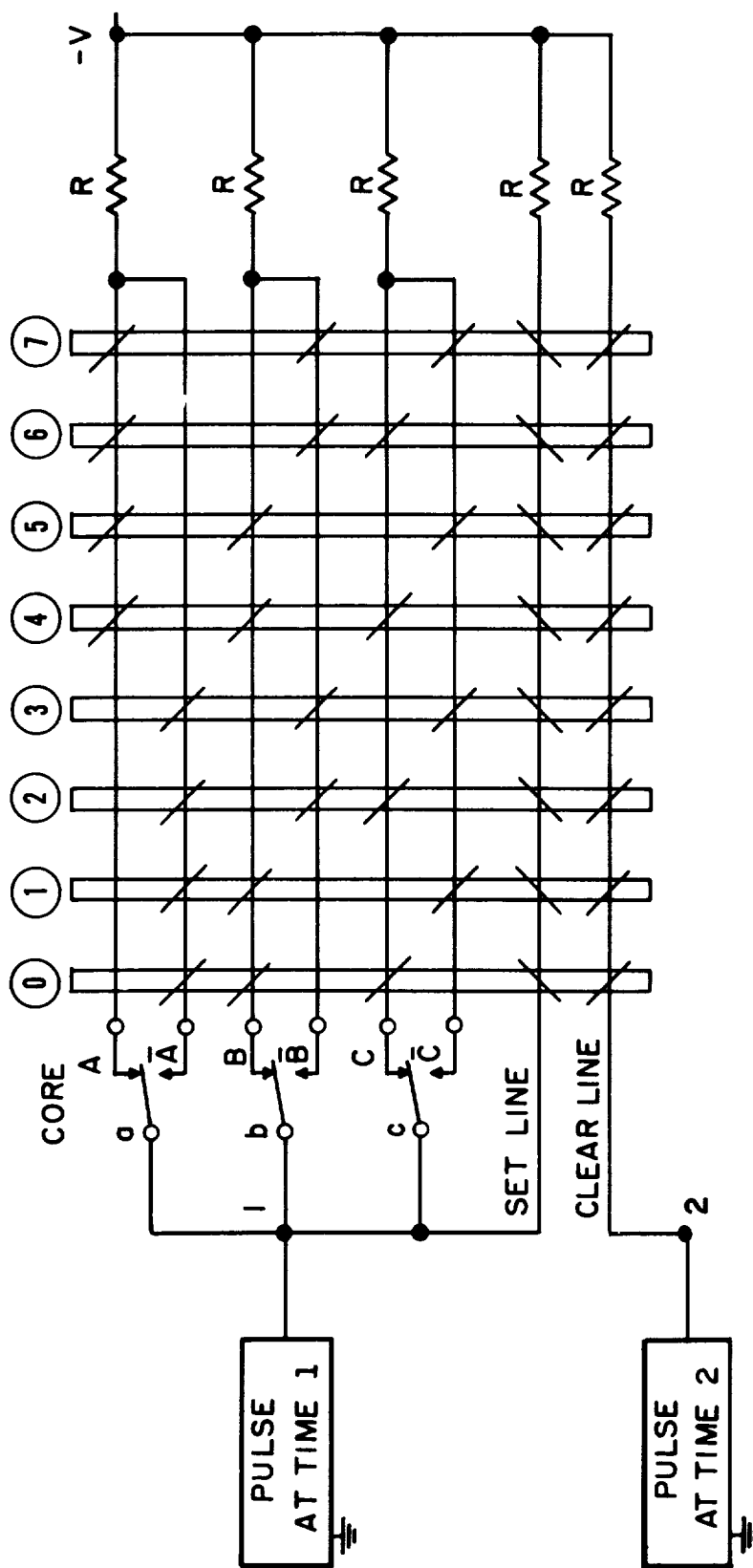


Fig. 1-2 Selection of a core in a core rope.

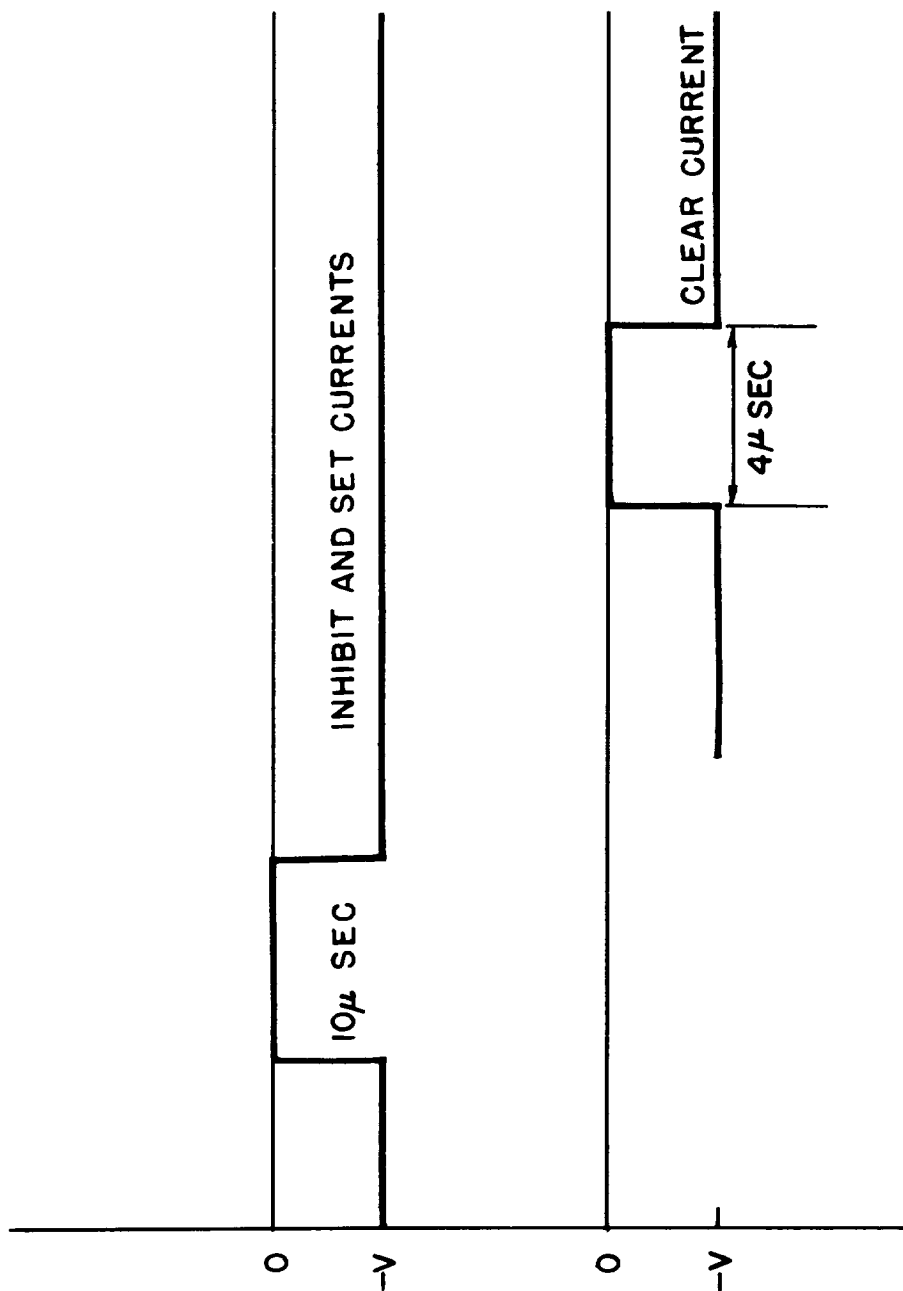


Fig. 1-3 Timing of inhibiting, set and clear currents.

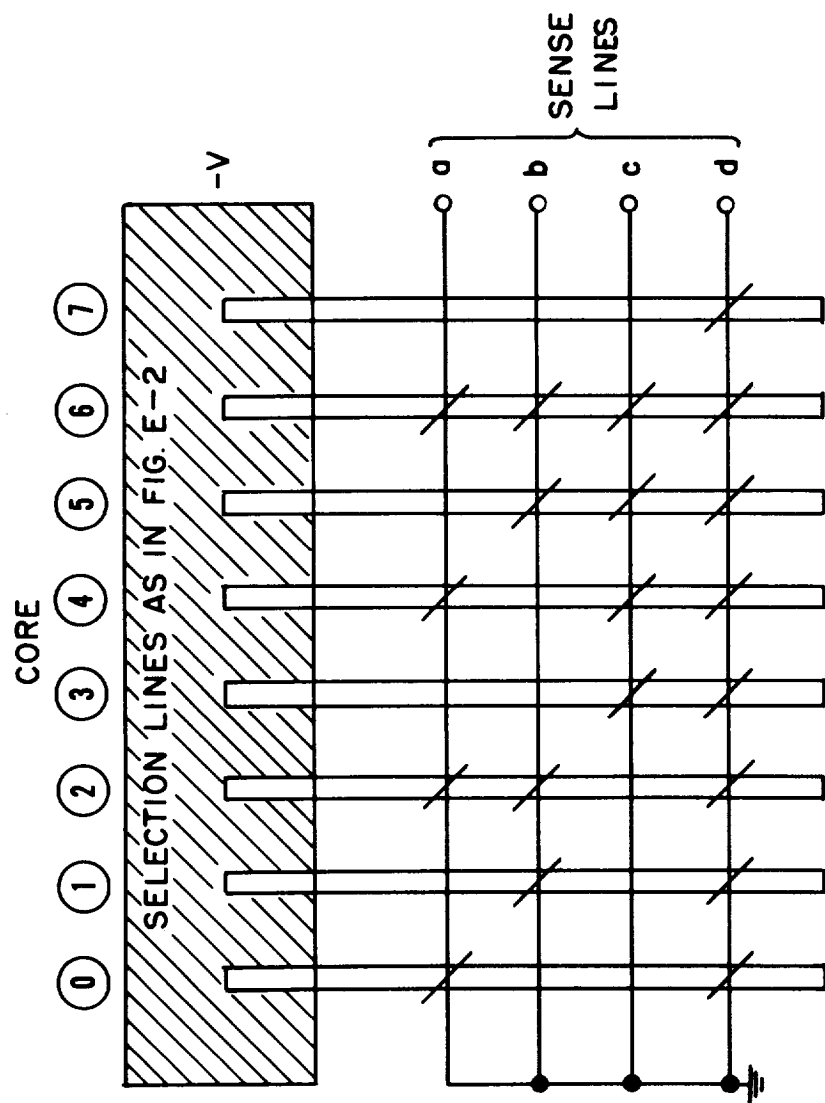


Fig. 1-4 Sensing wires on a core rope.

have an emf induced on them, and the remaining lines will not. Thus a pattern of ONE's and ZERO's in the resultant word can be specified by prescribing which lines do and do not thread the corresponding core.

In practice, the signals are observed at Clear time rather than at Set time, because the interference or noise induced by the currents in the many Inhibit lines masks the signals induced by the Set current. Nevertheless, the engineering problems of eliminating or minimizing the various sources of interference of this nature have provided one of the main obstacles that was necessary to overcome in the course of developing this computer. The situation is felt to be well under control for the memory sizes presently considered, but might constitute a problem in extending this technique to memories of significantly larger size.

Clearly, the size of the word stored per core is arbitrary, depending merely on the number of Sense lines. As a last remark, the whole core rope system may be thought of as a coding switch, with a correspondence between inputs (inhibiting currents) and outputs (emf's on Sense lines), rather than as a "memory" in the traditional computer sense. It is noted that the above method for selecting a core and making it switch is very economical, for it requires the logical minimum of $2N$ Inhibitors for selecting one out of 2^N cores.

B. Erasable Storage Registers

Each core of the rope just described may be regarded as a register, since it stores a word. However, that word cannot be erased and a new one put in its place because of the wired-in nature of the information, and hence these are known as "fixed registers". In order to store variables, it is necessary to have registers such that their contents may be changed; these will be called "erasable registers". The two kinds of memories will

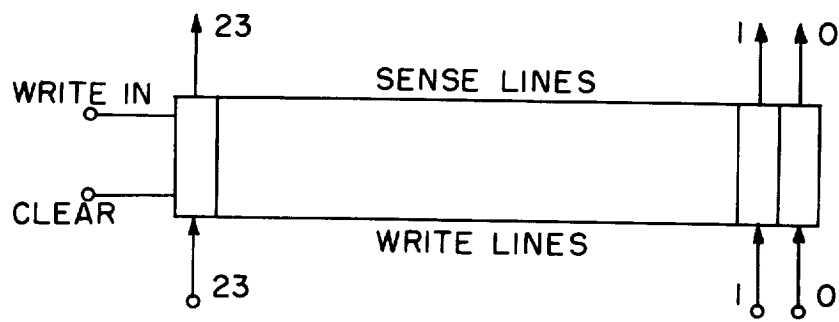


Fig. 1-5 Symbolic erasable register.

be referred to as the F memory and the E memory, or often simply as F or E. An operational description of an erasable register is shown in Fig. 1-5. Transfer of a word into and out of a register is done in parallel, i. e., at one pulse time. In an erasable register, a separate core is required for each binary digit to be stored, in contrast to the fixed register (rope core) which stores many digits per core.

Shown in Fig. 1-6 is a three-bit erasable register. When the control core is made to switch from a ZERO to ONE, transistor T1 saturates, and cores (0), (1), and (2) are cleared to ZERO. Those cores which are at ONE will induce an emf in the corresponding Sensing lines in the process of being cleared; ideally, those which are already at ZERO will not. Thus the contents of the register may be read out; e. g., to the special buffer register shown in the figure. Note that the process of reading out from an erasable register leaves all its cores at ZERO. If the original contents are to be preserved they must be written back into the register by a subsequent operation.

To write^{*} a word into an erasable register it is necessary to provide a current source on those writing lines for which a ONE is desired. This is indicated in idealized form by a set of relay contacts in the buffer in Fig. 1-6. Writing occurs when the control core is switched from ONE back to ZERO. At this time transistor T2 acts as a gate to permit current to flow from ground to the point marked (-V). A ONE is therefore placed in those cores which correspond to closed switches in the buffer register. The diodes of the Write lines prevent interactions between different registers.

A system of three erasable registers of three bits each,

*To write into a register is physically the same as to set all the register's cores; the convention used here is that whole registers are written into, while individual cores are set. Both are cleared.

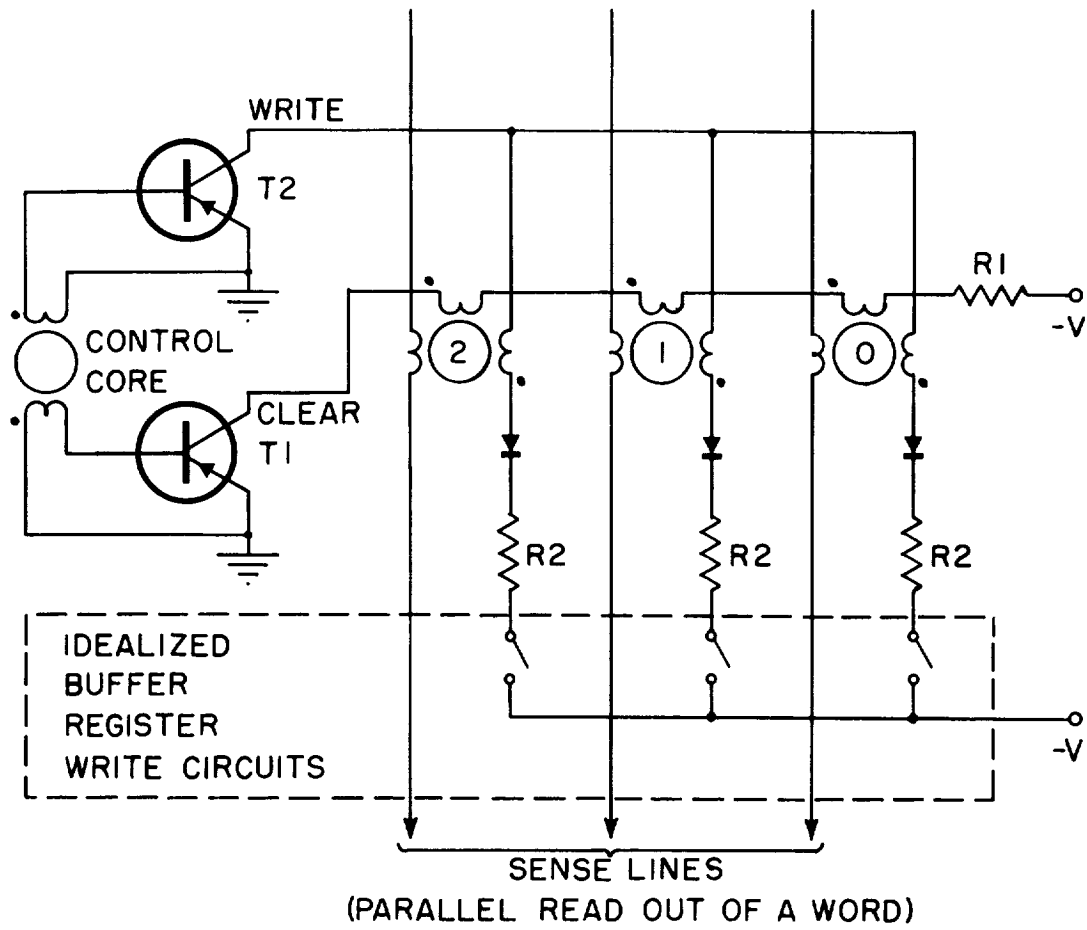


Fig. 1-6 Erasable register.

together with a buffer register, is shown in Fig. 1-7. It is seen that the buffer itself is simply another erasable register, separated from the other three by suitable sensing and writing amplifiers. In order to transfer a word from a register to the buffer, it is necessary to have time coincidence between a pulse on the Clear line of the register and a pulse on the Write line of the buffer. To transfer a word from buffer to register, the process must be reversed; i. e. , the Clear line of the buffer and the Write line of the register are pulsed at the same time. The active elements in the Clear and Write lines of each register in Fig. 1-7 are the respective transistors T1 and T2 in Fig. 1-6; thus each erasable register is composed of an array of cores and diodes plus two transistors.

The results of pulsing a register's Clear line without simultaneously pulsing the buffer's Write line will be to leave all the cores of the cleared register at ZERO, and the cores of the buffer unaffected. If the buffer is cleared and the register Write line is pulsed, without having previously cleared the register, then the contents of that register will be the "logical sum" or bit by bit meet of ONES from the word previously in that register (word A), and the word just cleared from the buffer (word B). In other words, the register will contain a ONE in those positions in which either word A or word B (or both) had a ONE.

Another possibility is that of transferring information between the buffer and two or more registers at the same time. If, at the time the buffer is cleared, the Write lines of the several registers are impulsed, the contents of the buffer will be transferred into each register (assuming them to have been cleared previously). Correspondingly, if two or more registers are cleared at once, the resultant information written into the buffer will be the logical sum of the contents of those registers.

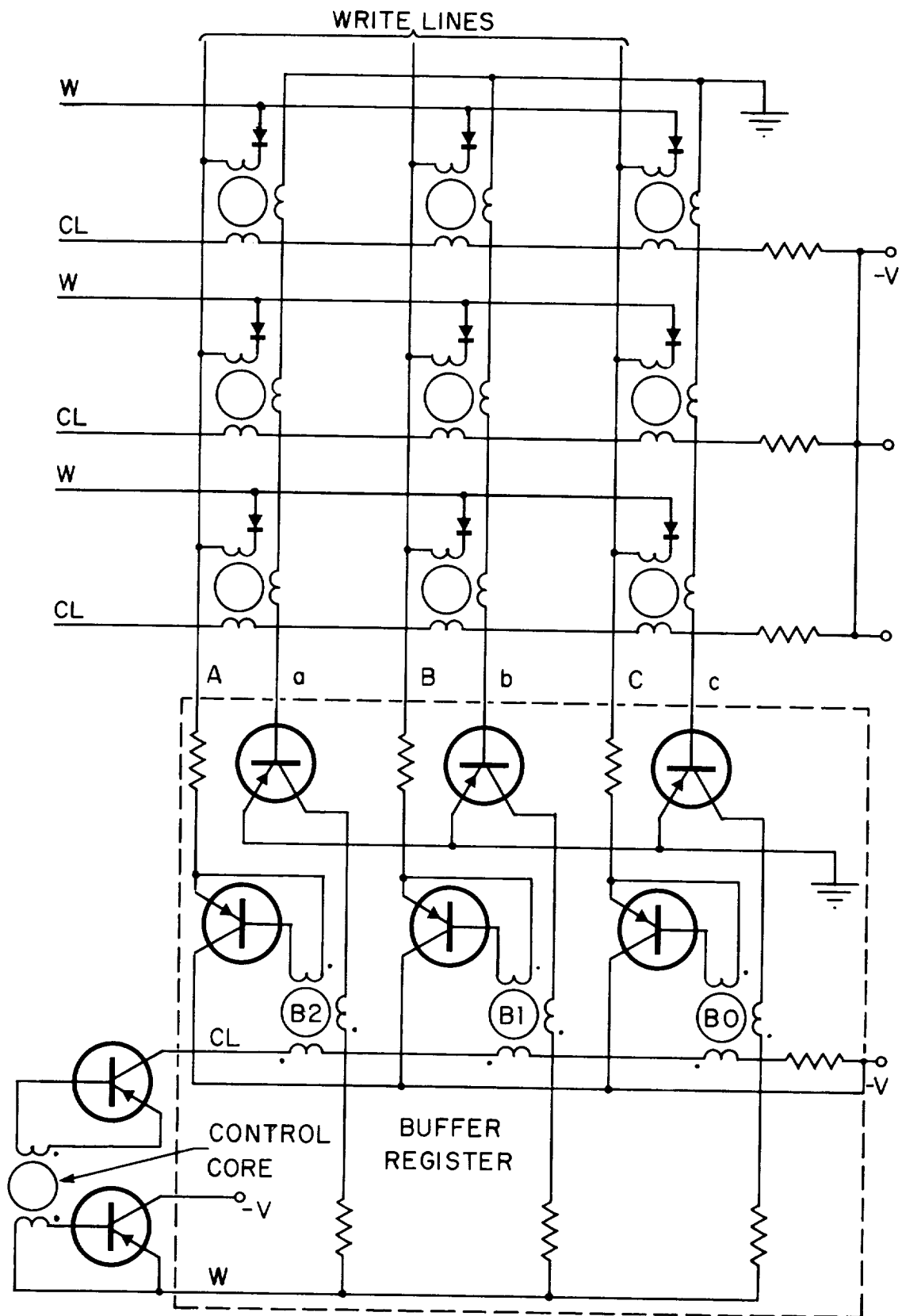


Fig. 1-7 Erasable storage system.

C. The Erasable Storage System

Shown in Fig. 1-8 is a ten-register erasable storage system, identical in operation to the small system of Fig. 1-7, but with additional useful notations illustrated. Examples are provided of several operations which can be performed on a word by use of a register with special rearrangements of the Write and Sense lines. For the sake of definiteness, the number of digits and the structure of the computer word are taken to be those for the representative machine. Each normal word has 24 binary digit positions numbered from 0 (low order bit) to 23 (high order bit). Digit 0 is customarily the parity bit (see Chap. 2, Section A) and, for numerical quantities, digit 23 is the sign bit.

The Write lines in Fig. 1-8 are represented by the arrow at the left, and directly over the arrowhead is, for each register, the list of Write lines (W) versus the bit positions of the register (R). Similarly the right hand numbers S and R refer to the Sense lines which thread the registers, and to their correspondence. For example, Register 1 has lines 23 through 0 connected to bit positions 23 through 0, on both the Write and Sense lines. This is of course the normal configuration; the wiring rearrangements in Registers 2 to 5 are the exception. Unless a register has markings to indicate the contrary, it should be assumed that it is a normal register (e. g. , Registers 6 through 10).

Registers 2 and 3 illustrate how wiring can be used to break a word into separate parts for manipulation by the computer. Transfer of a word from the buffer simultaneously into these two registers places digits 12-1 in Register 2 and digits 23-13 in Register 3. When Register 2 is subsequently transferred back into the buffer, only the low order digit portion of the original word will be involved; Sense lines 23-13 pass through no cores in Register 2, and will hence carry no pulses (i. e. , ZERO's) to the sensing amplifiers. The clearing of Register 3 will cause digits 23-13 of the original word to appear on Sense lines 11-1, with the

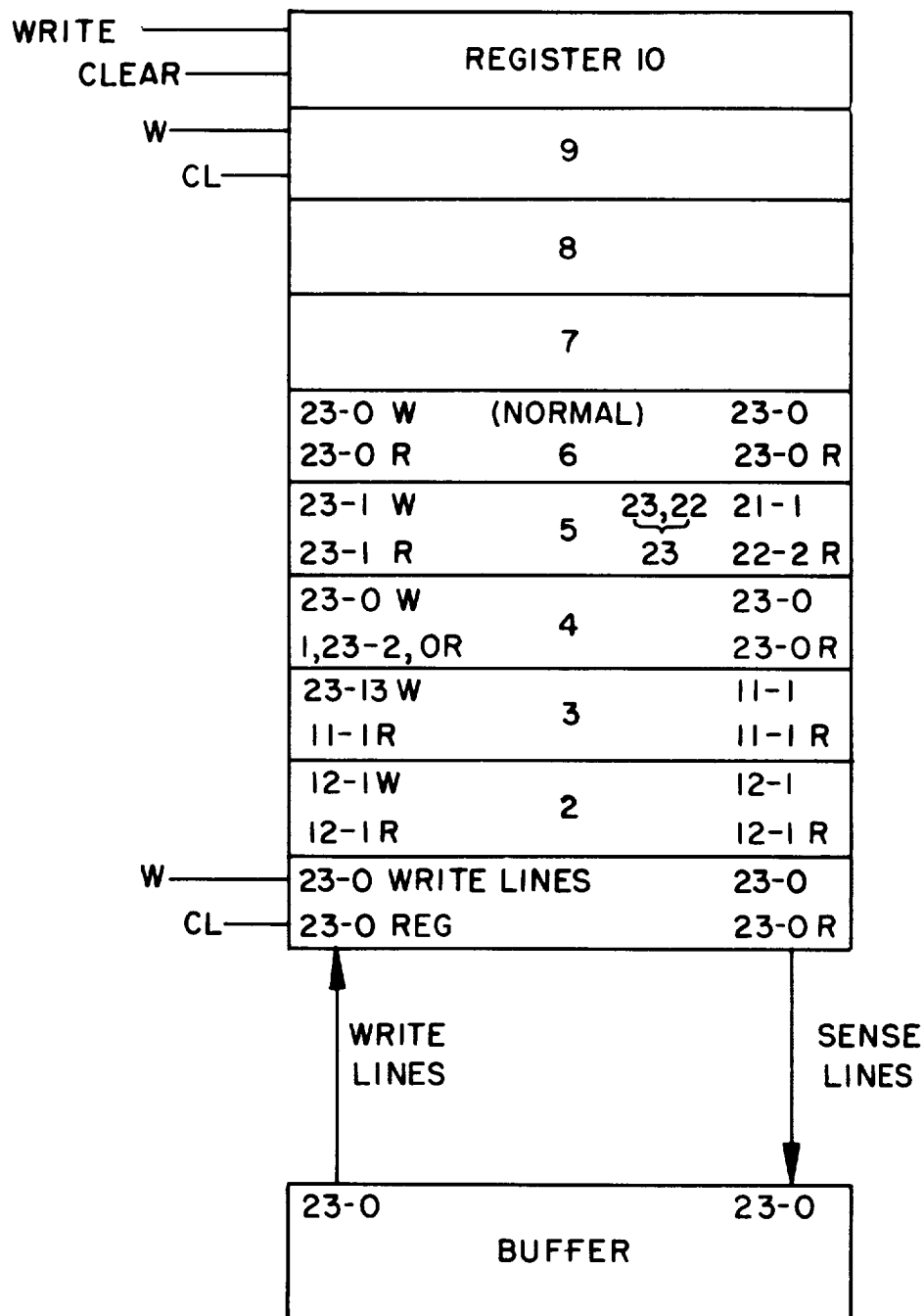


Fig. 1-8 System of ten erasable storage registers.

result that these digits are effectively shifted twelve positions to the right by the wiring of this register. Note that the parity bit (0) is omitted from the wiring of these registers, as it would lose its significance here.

As a further example of wiring rearrangements, Write line 23 is connected to bit position 1 in Register 4, Write lines 22-1 are connected to bit positions 23-2 respectively, and digit 0 is left invariant. The net result of writing a word into Register 4 is to shift it left one bit, with end-around carry, preserving the parity bit in the process. Further transformations may be obtained by manipulation of the Sense lines versus the register bit positions, as in Register 5. Here, Sense lines 23 and 22 both are impulsed by the clearing of core 23; Sense lines 21-1 are connected to bits 22-2. The result is to shift the word right by one digit, with the new digit position 22 being filled with a ZERO if the word is positive (digit 23 ZERO), or with a ONE if the word is negative.

The method for selecting a register is shown in Fig. 1-9. The figure shows a more complicated arrangement of the clearing and writing operations than is illustrated in Fig. 1-6. The added complexity provides greater flexibility and is further made necessary by timing considerations which are discussed in Chap. 1, Section D.

An erasable register is selected by means of a core rope, such as the one described in Chap. 1, Section A, where each core of the rope corresponds to a register. As before, all rope cores start at state ZERO. Selecting a core results in a single core (say core x of Fig. 1-9) being set to a ONE; nothing else happens at this time. When the rope core is reset by a pulse on the CL E line (for Clear E, and previously called the rope Clear line), core x is switched back to ZERO, which causes transistor T1 to fire. The firing of T1 in turn clears the erasable storage Register 100 (hence the name CL E), and sets core y to

ONE. All the cores in the same vertical column as core y (the writing cores of the E registers), started at ZERO; hence after $CL\ E$ only core y from among the writing cores is at ONE. The subsequent pulse on line $W\ E$ (for Write E) drives core y from ONE to ZERO, causing transistor T_2 to switch, and allowing the contents of the concurrently cleared buffer to be written into Register 100. As can be seen, writing into a register controlled in this fashion implies having cleared it previously.

D. The Complete Memory System

The memory system for this computer consists of fixed storage (F), erasable storage (E), and a special register (S) called the address register, by means of which storage is addressed. There are, in addition, other special registers on the erasable type called "central registers", which are used under direct control of the computer logic to execute the various instructions; the buffer B is an example*. These are discussed separately in Chap. 2, Section A; the "memory system" will be defined as the collection of registers which are controlled through the address register S .

The address register consists of two parts, denoted by the symbols S_d and S_c , which stand for the "direct" and "complement" parts of S . Each part is itself an erasable register of a length (12 bits, for the representative computer) sufficient for storing addresses of all computer words. The direct part of S stores the pertinent address, whereas S_c stores the ONE's complement of this address. The clearing of register S (i. e., the simultaneous clearing of S_d and S_c) by the control pulse $CL\ S$, causes the establishment of currents in the Inhibitor and Set lines (Fig. 1-2) required for switching a rope core associated

* To be quite accurate, S should properly be called a central register rather than part of the memory system.

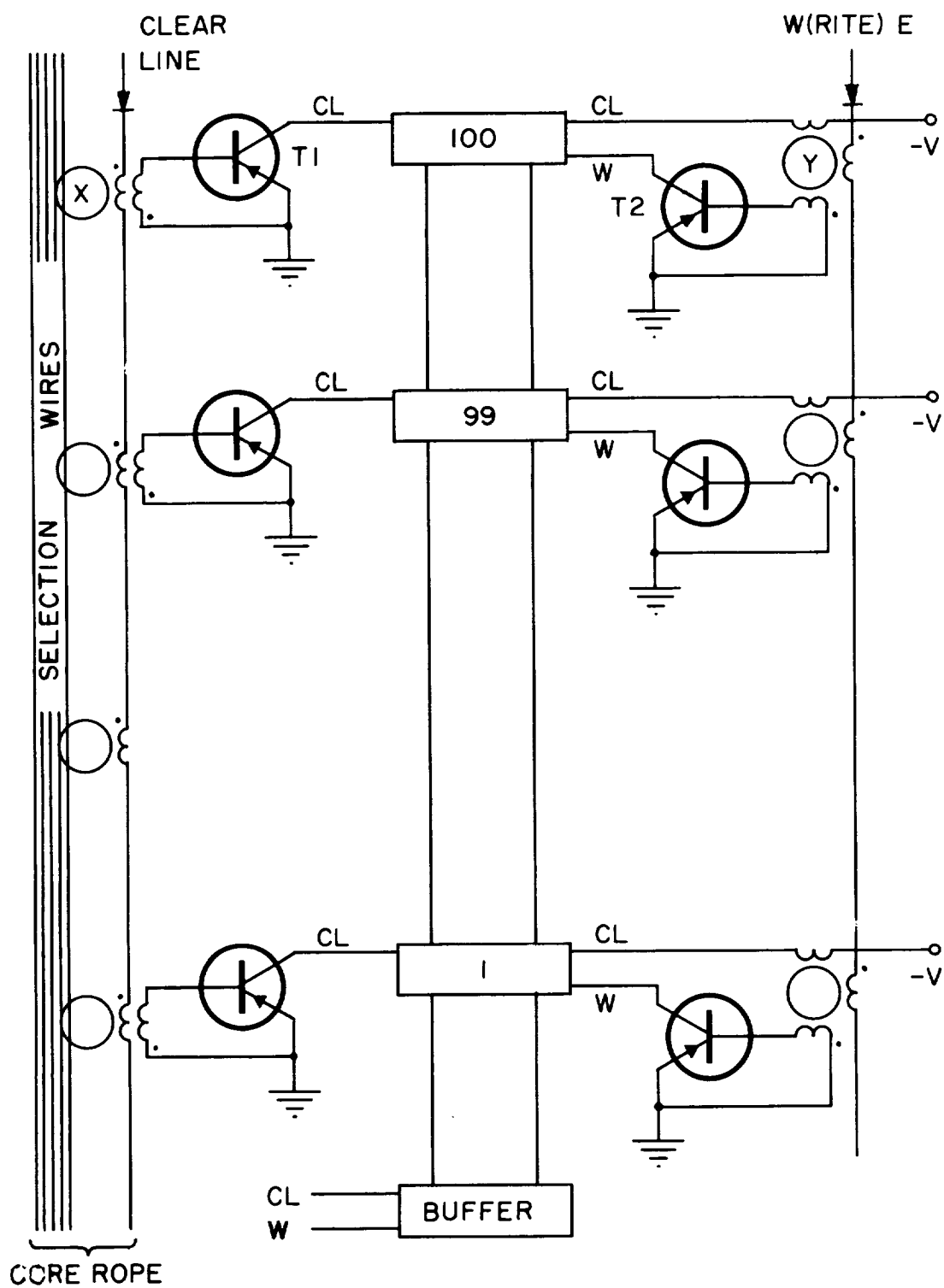


Fig. 1-9 Erasable storage selection system.

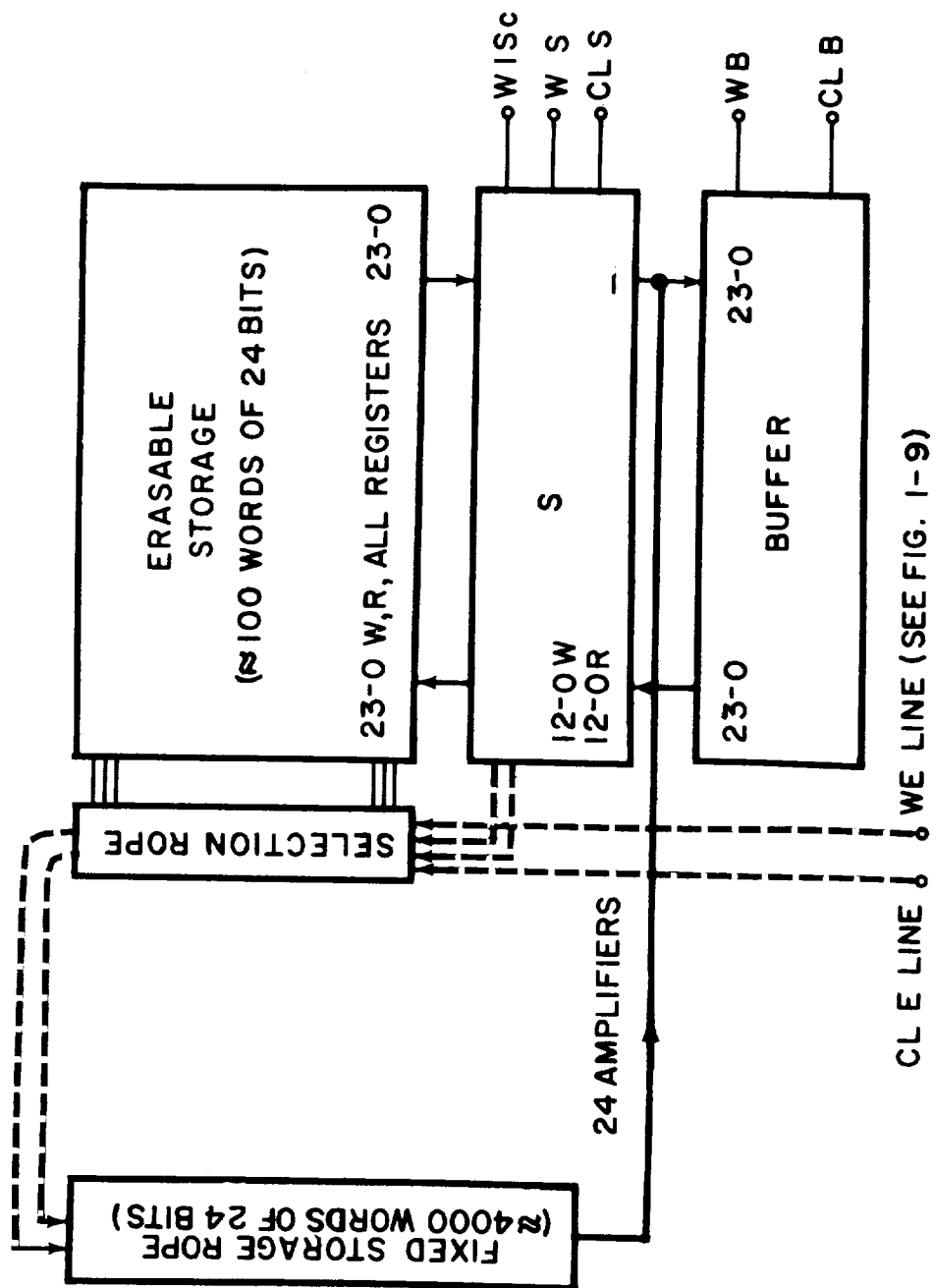


Fig. 1-10 Complete memory system.

with the address placed in S. These selecting currents are represented by the dotted lines of Fig. 1-10, and are seen to be common to the fixed storage rope and the erasable storage selection rope. The two ropes thus differ only in their purpose and name, and not in the method used for addressing or selecting a core.*

Each bit of Sd and Sc has a current generating circuit connected to it, which drives a corresponding rope Inhibit line. The control pulse CL S clears all cores of both Sd and Sc to ZERO, generating thereby a current pulse in each Inhibit line for which the corresponding core contained a ONE. The pulse CL S itself initiates the current through the Set line. Each pair of corresponding cores in Sd and Sc is associated with a pair of Inhibit lines. The fact that the contents of Sd and Sc are complements thus ensures that exactly one of each Inhibit pair is active.

Register S has several peculiarities. It is not connected to the Sense lines, and hence CL S does not result in the transfer of a word from S to the buffer. Furthermore, S requires a special control pulse WlSc, which results in all the cores in Sc being driven to ONE. This pulse is applied before transferring the address from the buffer to S, which is done by the pulse WS (and CL B, of course). The latter is the gating pulse which permits writing into both Sd and Sc. The Write lines are connected to Sd in the normal fashion for an erasable register. For Sc, however, the sense of the windings is reversed from the normal direction for writing, so that a current (i. e., a ONE on the Write line) causes the core to be driven from ONE to ZERO. Thus any digit position on the Write lines which contains a ONE

* By suitable rearrangements of the inhibit wires, it would be possible to have the addresses for erasable registers distributed amongst those for fixed storage in any manner whatever, if this were desirable, even though the erasable registers were themselves physically grouped together.

causes that digit in Sd to be set to ONE and in Sc to ZERO; conversely, a ZERO (no current) leaves Sd and Sc unchanged, i. e. , at ZERO and ONE.

It is now possible to list the sequence of pulses involved in reading from and writing into the complete storage system, and in so doing to explain the timing associated with Fig. 1-9. Each computer clock cycle is divided into two equal intervals called α time and β time. The computer runs on a two-beat system which, at α time, information flows into (or towards) the buffer and, at β time, towards erasable storage. Most control pulses have associated with them a specific time, either α or β , during which they may occur; they may not occur at the other time. For example, CL B can occur only at β , and WB at α . A more accurate statement would be that CL B may not occur at α , and WB may not occur at β , for it is not true that a pulse CL B occurs every β time or a pulse WB every α time.

Assuming an initial address to be in register S, then the following sequence of control pulses will result in a word being read out to the buffer from the register with that address, placed back where it came from if it came from erasable storage, and also placed in S. Each line shows the time, α or β , associated with these pulses.

<u>Time</u>	<u>Pulses</u>	<u>Comments</u>
α	CL S	
β	(wait)	No control pulse occurs
α	CL E, WB, WlSc	Reset rope, ONE into Sc
β	CL B, WS WE	

Although it may not seem so at first sight, this four -

pulse sequence takes care of both possible cases of reading from fixed or erasable storage. CL S selects a core in the rope regardless of whether it belongs to F or E, driving it to a ONE. CL E resets that core to a ZERO. If the core is associated with a register in E, then that register is cleared, its contents admitted to the buffer by virtue of the pulse WB, and the corresponding writing core (e. g. , core y of Fig. 1-9) is set to ONE. When the buffer is cleared at the next β time, the pulse WE resets the writing core to ZERO, permitting the contents of B to be written back into the E register from which it came. The word is also admitted into S by the pulse WS; it is noted that Sd and Sc have been set to ZERO and ONE by the successive pulses CL S and WlSc prior to the pulse WS.

If the core switched by CL S belongs to F, then the subsequent CL E and WB causes a wired-in word to be read into the buffer through a set of amplifiers* (Fig. 1-10). The last row of pulses again causes the buffer to be cleared and S written into. WE here has no effect, for no writing core has been set to a ONE; hence none of the E registers is affected.

The sequence for replacing the word in an E register by a word previously stored in B is very similar, namely:

α : CL S
 β : (wait)
 α : CL E, WlSc
 β : CL B, WS, WE

The only difference is the absence of WB from the second α pulse. If this pulse is absent, the word cleared from the E

* The main electrical distinction between F and E signals is that those from F are weak and occur in the midst of noise, while signals from E are strong and free from interference. Hence the amplifiers.

register is lost. The pulses occurring at the following β time will cause the word previously placed in B to be transferred into the E register. It was assumed that the address originally in S was an E address, but the wired-in nature of F is such that nothing would happen if the address were an F address.

The control pulse sequences illustrated in this section accomplish quite elementary and restricted objectives, and are introduced primarily to discuss principles of communication with the computer memory. In the next chapter, after a discussion of central registers, consideration is given to the formation of control pulse sequences into computer orders in the accepted sense. In particular, an order code is given in more or less complete form for the representative computer, and possible variants are also presented.

CHAPTER 2

EXECUTION OF COMPUTER INSTRUCTIONS

A. Central Registers and Overall Organization

The overall organization of the computer, shown in Fig. 2-1, includes storage, discussed in Chapter 1, input-output equipment, discussed in the next chapter, together with the sequence generator and central registers which comprise the main logic for executing computer instructions. In this chapter the central registers and sequence generator are discussed, first in terms of generalities, and subsequently in somewhat more specific form in relation to the order code for the representative computer. In spite of the latter discussion, however, it should be emphasized that we are here concerned with a diverse class of possible machines rather than primarily with the specific computer selected as "representative"; this should be evident from the number of alternatives which are presented from time to time.

Central registers are, by definition, all those registers which are not addressed by means of the address register S, but whose control pulses are provided directly by the sequence generator. Registers S and B, for example, have previously been described as central registers; the pulses CL B, WS, etc., are provided by the sequence generator, as are the others listed in Chap. 1 Section D as necessary for reading from memory. The outputs from the sequence generator are, at pulse times α and β , pulses which bear names like CL B, CL E, WE, WS, WSc, etc.

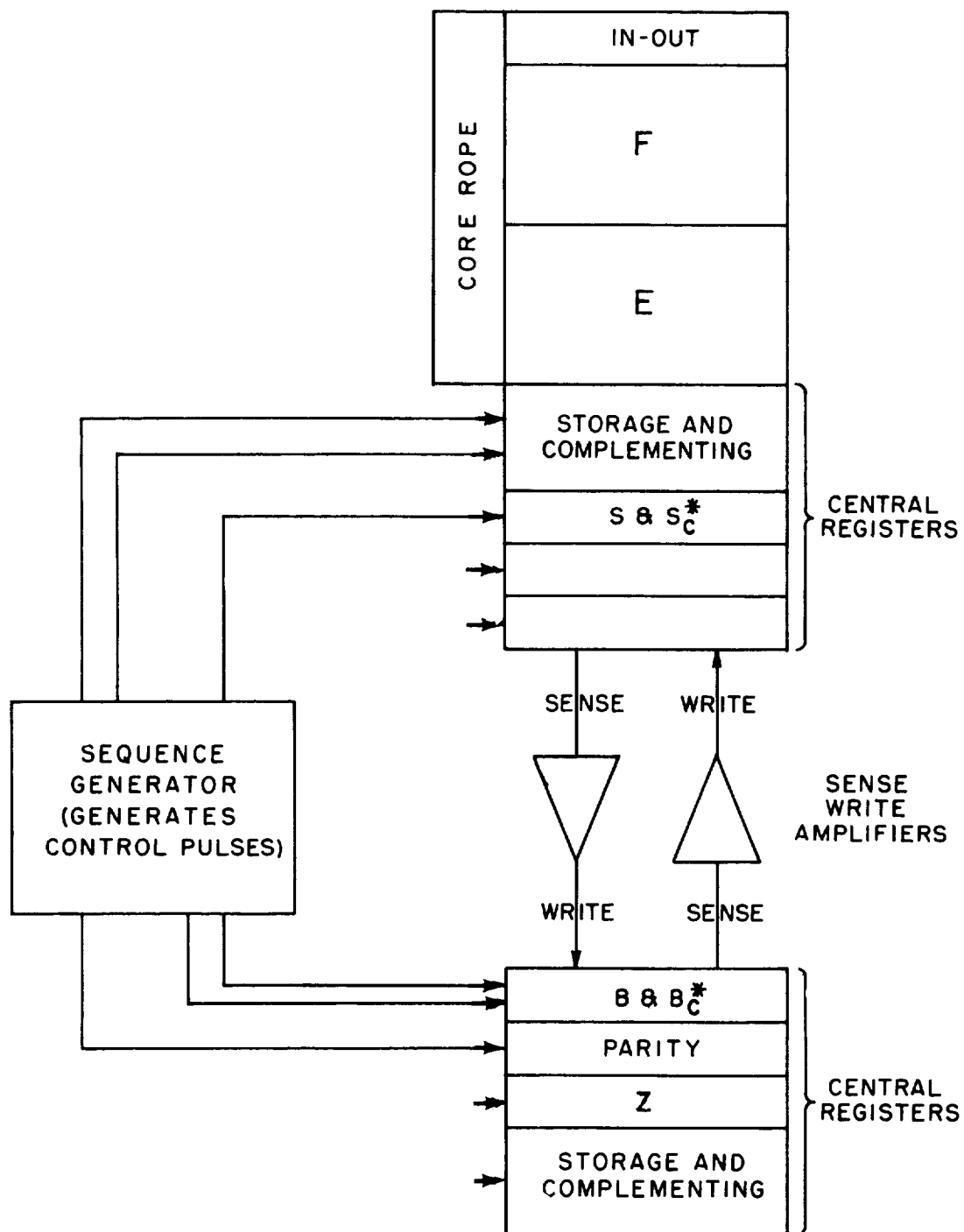


Fig. 2-1 Over-all organization of the representative computer.

Most of these control pulses go to the central registers; however, a few like CL E have other purposes. What follows is a brief description of the various kinds of central registers and a possible central register organization.

The overall computer organization is shown in Fig. 2-1. There are two sets of Sense and Write lines, one for the transfer of words from the core rope side to the B side, and one for the transfer of words from the B side to the rope side. Central registers may be on either side of the Sense-Write amplifiers and are of three kinds: special-purpose registers which have extra circuitry, complementing registers, and ordinary storage registers. The special-purpose registers are the address register S discussed previously, the Z register, the parity register P, and the buffer B. These are discussed separately in subsequent paragraphs. As in E storage, ordinary storage registers may occur with rearranged wiring to secure various special effects.

The "rope side" of the Sense-Write amplifiers is henceforth called the α side of the register system; i. e., the side on which registers are cleared at α time. The "B side" is correspondingly called the β side. The division of the computer into an α side and a β side, together with the electrical properties of the erasable storage scheme, helps explain the working of the computer on a two-beat system and why each control pulse has a particular time (α or β) at which it occurs.

At α time information flows from the α side to the β side and at β time from the β side back toward the α side. It would appear that both actions could take place simultaneously, because separate sets of Sense-Write amplifiers are used. Consider, however, the signals from the sensing windings on the α side. These result from certain cores on the α side being cleared from ONE to ZERO at α time. If other cores on the α side are being driven from ZERO to ONE by a simultaneous writing operation, a reverse emf is induced in the corresponding sense lines

which cancels the desired signals where overlapping occurs. As a result, there is a significant coupling from the write side of one set of amplifiers to the sense side of the other set which prevents them from being active at the same time. Various electrical schemes could eliminate or avoid this restriction; however, the present arrangement appears to offer a simple and reliable system.

The Z register is a special-purpose register that adds "one" to a number read into it, so that a transfer into and out of Z converts the number N into $N + 1$. It is used, on the one hand, to store and increment the address of successive computer instructions when these occur in sequence. On the other hand, it is used on a time-shared basis to increment the various program or input counter registers (see Chap. 2, Section C and Chap. 3, Section E respectively). The electrical nature of the Z register is fairly well established; however, at the moment of writing, there are one or more unsettled questions concerning its logical integration with the rest of the computer. These relate to the treatment of input situations in which pulses to be counted may be either plus or minus, and to the length of the Z register.

The parity register, P, accepts inputs on all 24 Sense lines and computes whether the number of ONE's in the word being read into it is even or odd. Its output is one bit long and occupies digit position 0; if the number of ONE's in the word is even, then the parity bit will hold a ONE. A parity system is employed in which the total number of ONE's in a word including its parity bit, must be odd to pass the parity test. When such a word is read into P, the contents of the parity bit will be zero, so that an alarm will not be triggered when P is subsequently cleared and tested. If, on the other hand, a 23-digit word is read into P with its parity bit missing, the parity register computes and stores the correct value in digit position 0. Thus the parity may be transferred to E storage along with digits 1-23 by

clearing P at the same time the remainder of the word is stored.

The buffer register B is a special central register used as an ordinary register in communicating with storage and also used in the decoding of computer instructions. It is composed of two parts, Bd and Bc, as is the address register S. The direct part, Bd, is a full-length register; however, the complement part, Bc, occupies only the digit positions used by the order code (digits 13-18 for the representative computer). The latter bit positions of both Bd and Bc are wired to generate Inhibit pulses in a special short core rope, for selecting a sequence in the sequence generator. They operate in the same way as Sd and Sc, being cleared by the special pulse CL Bc. The pulse called CL B actually should be entitled "CL Bd" since it leaves the complement portion of B untouched. Also, the pulse CL Bc clears corresponding digits in both Bc and Bd and further provides a gate to permit the Inhibit currents to flow. The latter is necessary since Bd is used for other purposes.

Central storage registers are identical to ordinary erasable storage registers. They each require two control pulses, CL register and W register (for "Clear" and "Write"). Complementing registers are similar to storage registers but with one extra common winding added. This winding is used to set all the cores in the register to ONE, and is pulsed by a control pulse W1 Register (for "Write Ones"). The polarity of the windings which connect each core of a complementing register to a Write line is reversed relative to the polarity of a storage register. Thus if a word is written into a complementing register which has been set to all ONE's, the register will then hold the bit by bit complement of the word.

The number and arrangement of central registers of the storage and complementing types is a function of the order code desired, and of the way in which the order is translated into a sequence to be generated by the sequence generator. In

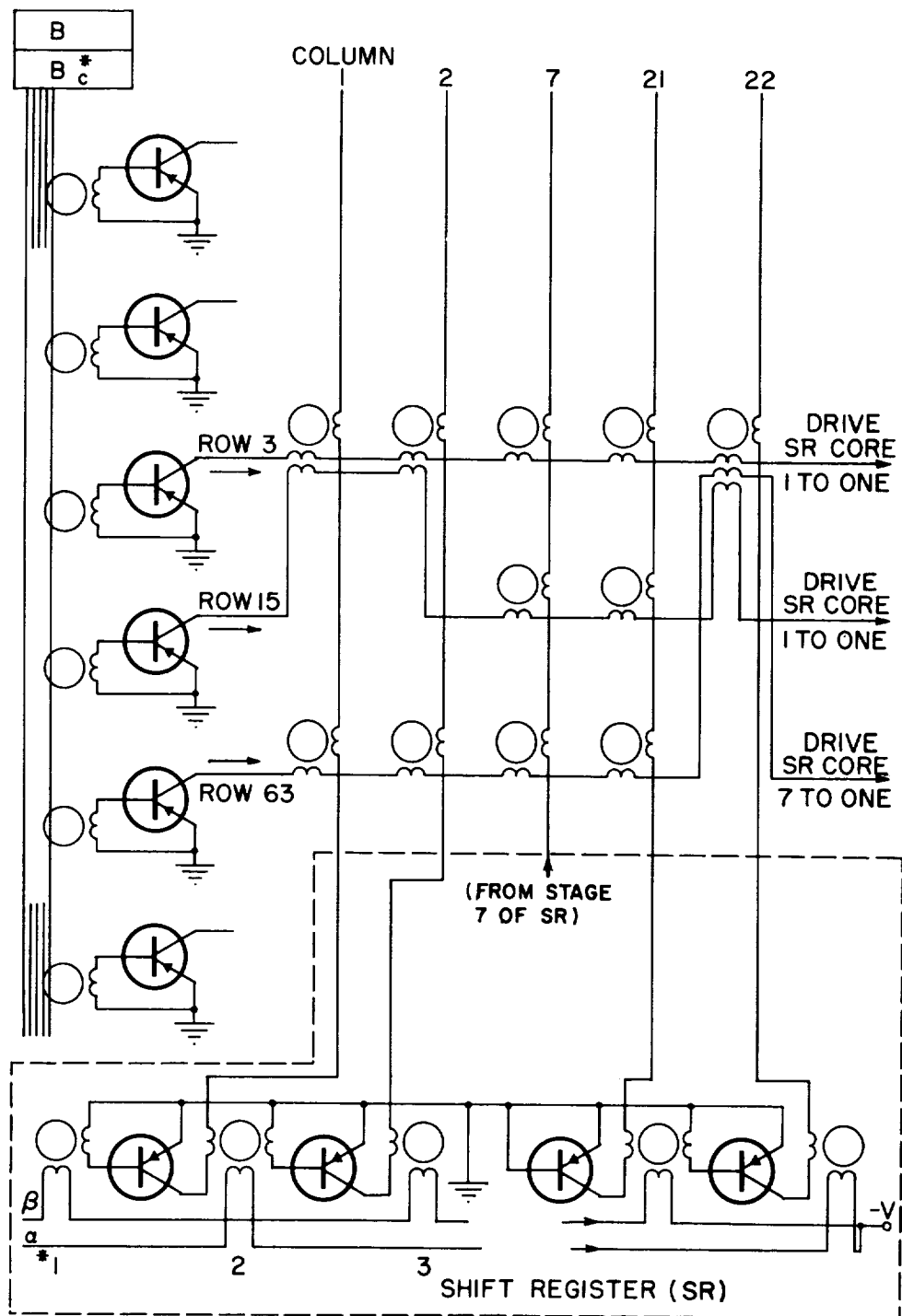


Fig. 2-2 Sequence generator.

general, the more powerful the order code, the larger the number of such central registers. These registers are in themselves no more expensive than those in erasable storage, but addressing them must be done through the sequence generator which is a costlier process than addressing through a rope. Later in this report specific arrangements of central-registers will be discussed in order to illustrate how certain things can be done and what alternatives there might be. This discussion is necessarily tied to a discussion of the order code and of the sequence generator. The address register S, the "add one" register Z, the parity register P, and the buffer B are common to all computers under discussion.

B. The Sequence Generator

The sequence generator forms what is commonly called the "logic" of a computer. Other common names for it are program counter and microprogram store. The basic idea is that of associating a sequence of control pulses with each kind of order stored in F or in E. The sequence involved in transferring the contents of some register X into X and central register A, which bears the name "clear and add X", or CA X, is one such order. That involved in storing the contents of A in X, called "transfer to storage X", or TS X, is another. By means of the sequence selection circuits (mentioned above in connection with the buffer), each combination of bits in the order code digit positions of an instruction is made to correspond to the particular pulse sequence for execution of that instruction. For reasons which will become clear later, all sequences produced in the sequence generator are called "zero level sequences".

The order structure of the representative computer is such that bit 0 (numbering from low-toward high-order digits) indicates the parity of the word (cf. Chap. 2, Section A), and the next twelve bits (bits 1-12) are for any address which forms

part of the instruction (e. g., the "X" in CA X). Digits 13 to 18 form the symbol or code for an order, i. e., the "name" of a sequence. The remaining five digits may have other optional uses which, in effect, cause them to act like part of the order code.

The sequence generator is a complicated device, and hence a simplified case will be discussed first. Consider a rectangular matrix, made of cores, with 64 rows and 22 columns (Fig. 2-2, 2-3). At the bottom of each column there is a core - transistor circuit, symbolized in Fig. 2-2 by a square box with a number inside it, which, when fired, drives all cores of that column to ZERO and drives to a ONE the core in the next transistor-core element. These 22 elements form a shift register. All the cores in the matrix start at ZERO. Selecting a sequence means setting a row of cores to ONE's and placing a ONE in the first position of the shift register (Fig. 2-3). Hence only one core of each column may be switched to ZERO by the corresponding shift register element*. As the shifting pulses α and β (Fig. 2-2 and 2-3) shift the single ONE in the shift register through all positions from (1) to (22), all the matrix cores set to ONE (forming a row) are cleared to ZERO one by one. Each control pulse has a wire associated with it, which threads (or "not threads") the cores of each row. Thus when a row of cores switches in sequence, the wire has a pattern of outputs and "no outputs" impressed on it.

For example, consider control pulse A (Fig. 2-2) if the selected row is (2). The only output (i. e., the only control pulse) will occur at shift time 7, i. e., when the ONE travelling down the shift register is being transferred from position (7) to position (8). If the sequence selected had been sequence (= row) (3), then there would have been outputs at time 1 and 2. Notice further

* All cores of that column are driven to ZERO, but only one switches, since only one of the cores of that column was switched to ONE in the first place.

that sequence (1) involves neither pulses A nor B, while sequence (63) involves both.

There are several variants of the matrix just described which enhance the flexibility of the sequence generator and permit certain economies. The particular use of these variants depends directly on the properties of the sequences, which will be discussed later, and hence the variants are presented here as "tricks" which can be used if certain situations arise.

The most obvious variant is illustrated in Fig. 2-3 and consists in modifying the meaning of "row". Selection of a sequence means setting to a ONE some cores of the matrix, where each column will have at most one core set to ONE. Many sequences may have common parts; for example, they must end with an identical set of control pulses, for calling forth the next instruction. Then the same core in the matrix may be set by the various sequences, as in the case of the core of row (1), column (22), Fig. 2-3. In fact, that core is the only core present in column (22). Also, it is possible that a sequence may require a waiting period of one or more pulses, in which case the sequence selecting wire which sets a "row" to all ONE's will simply not set any core of the pertinent columns to a ONE. For example, sequence (15) (row 15) does not set any core of column (21).

A further "trick" consists of the use of the sequence selecting mechanism for specifying the starting point in the shift register. In Fig. 2-3, sequences (3) and (15) start at shift position (= time) 1, while sequence (63) starts at time 7. This scheme is useful in avoiding unnecessary waits and delays. Thus, the sequence CA X (clear A and add the contents of register X into A) and the sequence CA* X, in which X is now interpreted as the location of the address of the data, are such that the set of pulses for CA X is a subset of those for CA* X. CA X starts at time 7; CA* X starts at time 1 and is identical to CA X from time 7 on. The wire which selects sequence CA X drives core 7 of the shift

register to ONE, thus avoiding a wait of 6 pulse times.

A third and final variant of the basic scheme consists of providing for branches. This is done by using a sequence to test the state of a core, e. g. , the core in which the sign of the quantity presently in A is stored. This result is then used to select one of two sequences, i. e. , to set to ONE's the cores of one of two "rows". This is discussed further in the next section.

C. Representative Order Code System

For purpose of illustrating the variety of computer operations which can be accomplished by rather simple wiring and logic, a representative system of central registers and instructions, is described in this section. The system is not completely described here; details are not sufficiently settled to make this possible. As show below, however, some relatively sophisticated operations can easily be achieved.

The central register system shown in Fig. 2-4 contains registers S (Sd, Sc), B and Bc, P, and Z which have already been discussed. In addition are shown ordinary register D, registers A and Q, each of which consist of a direct and complement part, complementing registers R1c and R2c, together with register CYc which simultaneously complements a word and cycles it left one digit with end-around carry. Four control pulses affect register A: WAd, and WAc for writing respectively in the direct and complement portions, CL A, which clears both halves simultaneously, and W1Ac to place ONE's in Ac. Register Q possesses corresponding properties. Registers R1c, R2c, and CYc, possess three control pulses each for writing, clearing, and placing ONE's in the cores before writing.

To illustrate an entire sequence of steps comprising one relatively simple order, consider the instruction "Clear and Add X", CA X, as described previously. Control pulses for CA X are given in Table 2-1. Each instruction must bear the responsibility for calling forth from storage the next instruction in

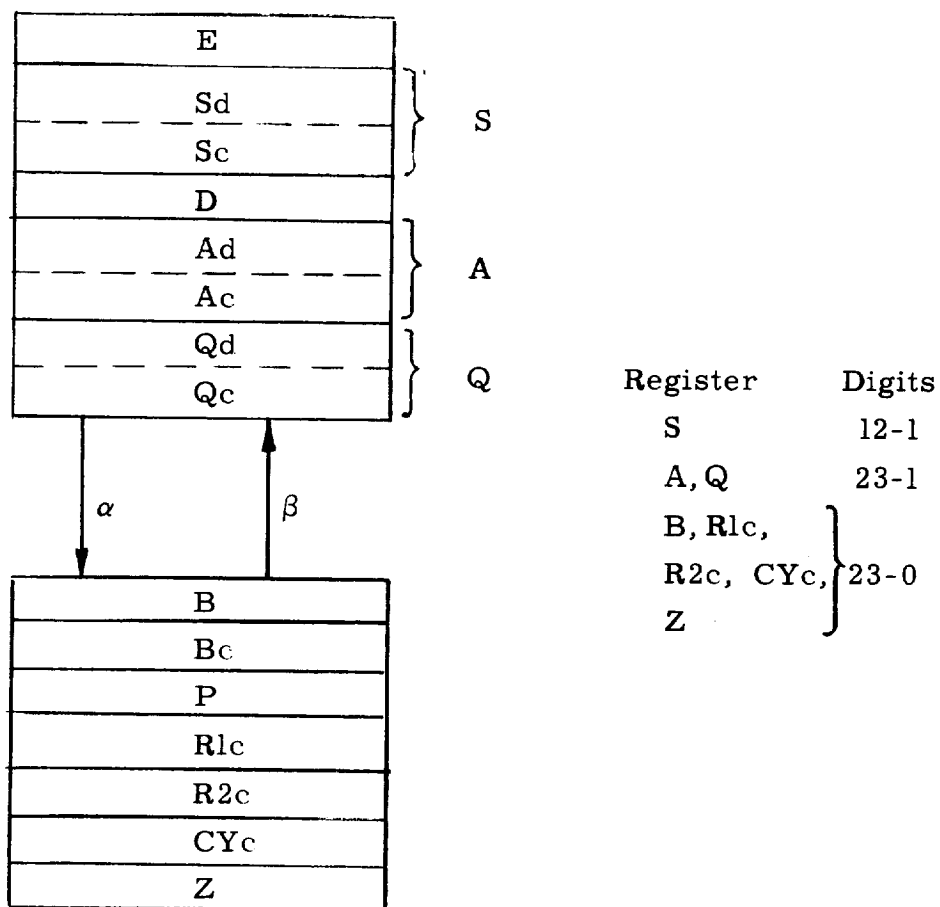


Fig. 2-4 Representative Central Register System

TABLE 2-1

CA X, Clear and Add X, C(X) into A, 8 pulses

- 1 α CL S, CL A, CL D
- 2 β (no action)
- 3 α CL E, WB, WP, W1Sc
- 4 β CL Z, WS, WD
- 5 α CL S, CL D, WZ, W1R2c
- 6 β CL B, CL P, TP, WAd, WE, W1Bc
- 7 α CL E, WB, WP, WR2c, W1Sc
- 8 β CL B, WE, WS, WD, CL Bc, CL P, TP

sequence, after its own action is complete, so that the control pulses form a never ending sequence. A convention which states at what point of the process the "new" order takes over the "old" must be adopted here. The operation CL Bc, shown in the last line of Table 2-1, sets the appropriate line of cores in the sequence generator core matrix for the new instruction; it is therefore reasonable to define the α - pulse time which immediately follows as the first pulse of the next order.

To follow the actions described by the control pulses in Table 2-1, it is necessary to know that certain of the central registers are assumed to have been left with particular contents by the previous order. Line 8 in the table is the last line executed by every instruction; at the time of its execution, the next order is in B. By the action of line 8, the order is written back into E (if it came from E), into D, and the address portion is written in S. Thus at the time of initiating CA X, in line 1 of the table, the last line of the previous instruction has placed the address X in S and the entire order in D. The effect of line 1 for CA X is to clear D, to clear the accumulator A in preparation for later receiving the contents of X, and to initiate read-out of address X through setting the corresponding rope core.

Line 3 of Table 2-1 transfers the contents of X into the buffer, B, and writes them also into the parity register P. The Z register at this point contains the address of the next instruction, $N + 1$; line 4 transfers this address to S as the initial step in the action of calling the next instruction forth. The number $n + 1$ is also saved in register D, from where it is cleared in line 5 back into Z. The result leaves $N + 2$ in Z as the address of the instruction after next. Neither of the actions CL S or WlR2c interfere with this transfer since S has no sense wires and since the sense wires for R2c are in use only at β time. The control pulses WlR2c in line 5 and WR2c in line 7 serve the purpose of saving in R2c the complement of each instruction

executed; this feature is required only in a very special situation to be discussed later. It is noted that W1R2c prepares R2c to receive a new word just as effectively as clearing an ordinary register. All cores are left at ONE's, regardless of their previous status.

In line 6 the pair of pulses CL B, CL P places on sense lines 23-0 the full word, including its parity bit, that was read from storage. The parity bit comes from digit 0 of B in this instance; the contents of the single bit called register P are properly a ZERO at this time since an entire 24 digit word from storage, with presumably correct parity, was read into P in line 3. Thus the pulse on digit 0 sense line is the logical sum of a ZERO and digit 0 of B. The control pulse labelled TP (Test Parity) activates a test to verify that P does indeed contain a ZERO as it is cleared. An alarm, as yet specified, or an error diagnosis procedure is to be initiated if this is not the case.

Line 6 writes the original word from storage back into E if required, and writes digits 23-1 of the word into part Ad of the accumulator A. Thus the contents of register X have been placed in A, and X is left unaltered. The next instruction is already in the process of being called forth by the pulse CL S of line 5 which sets the appropriate rope core. The occurrence of the pulse WE in line 6 after the pulse CL S in line 5 may offer some confusion, until it is recalled that the write core for an erasable register is set only when the rope core is reset. Thus the rope core for the next order, i. e. , core N + 1, may be set prior to the time at which the erasable storage write core (if any) associated with register X is reset.

The pulse W1Bc places ONE's in Bc to enable the complement of the order code to be obtained through the WB pulse in line 7. The fact that W1Bc occurs simultaneously with CL B would at first glance appear to violate basic principles. Register

Bc contains no sense windings, however, except those connected to the logic decoding rope. Further, CL B leaves Bc unaffected; only WB is common to B and Bc.

The principal action of line 7 in Table 2-1 is to call forth the next instruction from storage, writing it directly in the buffer B (and in Bc, as well), in the parity register, and in the form of its complement into R2c. Line 8 has already been discussed briefly above; the new order is written back into E if relevant, into D, its address portion into S, its parity is checked, and the pulse CL Bc transforms its order code into appropriate action in the sequence generator.

Considerable detail has been brought into the analysis of the pulse by pulse action of a single elementary order, CA X, because in so doing a considerable portion of the logic of most of the rest of the orders is also presented. In the following paragraphs and tables other typical orders are presented, generally with no more than a few sentences of description attached to each. These descriptions are made somewhat more concise by use of a few simple abbreviations and conventions. The order being executed is understood to have been found in register N, and the next order is located in register N + 1 unless otherwise noted. The address portion of the order is denoted by X, as in CA X above. The contents of any register are denoted by the symbol C() with the symbol for the register inserted in the parentheses; e. g., the contents of A are written C(A). The system of arithmetic used here is that in which the negative of a word is simply its ONE's complement, denoted by following that word with a prime: - C(X) is the same as C(X)'. Finally, a register X, A, or Q is understood to be left unchanged by the action of an order, unless a change is indicated specifically.

The order TS X, to store the contents of A, has been mentioned previously. As seen from Table 2-2, it differs from CA X only in that the contents of A are transferred to B and P

in line 1, instead of being discarded, and the contents of X are discarded in line 3 instead of being transferred to B and P. Note that CL P in line 6 is here not accompanied by the parity test pulse TP as it was before. Instead, the correct parity bit for the word read into P in line 1, on digits 23-1, is supplied as the output of P on digit 0. The physical action of the parity circuit is the same in the two cases, but its function is different. Digit 0 of B is ZERO in line 6, since only positions 23-1 were filled from A in line 1.

TABLE 2-2

TS X, Transfer to Storage X, C(A) into X, 8 Pulses

- 1 α CL S, CL A, WB, WP
- 2 β (no action)
- 3 α CL E, CL D, W1Sc
- 4 β CL Z, WS, WD
- 5 α CL S, CL D, WZ, W1R2c
- 6 β CL B, CL P, WAd, WE, W1Bc
- 7 α CL E, WB, WP, WR2c, W1Sc
- 8 β CL B, WE, WS, WD, CL Bc, CL P, TP

The instruction CS X, to clear A and place the negative of C(X) in A, is seen in Table 2-3 to differ from CA X only in the pulse W1Ac in line 2 and in the replacement of WAd by WAc in line 6.

TABLE 2-3

CS X, Clear and Subtract X, C(X) into A, 8 Pulses

- 1 α CL S, CL A, CL D
- 2 β W1Ac
- 3 α CL E, WB, WP, W1Sc
- 4 β CL Z, WS, WD
- 5 α CL S, CL D, WZ, W1R2c
- 6 β CL B, CL P, TP, WAc, WE, W1Bc
- 7 α CL E, WB, WP, WR2c, W1Sc
- 8 β CL B, WE, WS, WD, CL Bc, CL P, TP

If, the first line of control pulses for CA X, the pulse CL A is omitted, the result is an order LA X as described in Table 2-4, called "Logical Add X". The effect of this order is to place in A the word which is the union or logical sum of the previous C(A) with C(X); i. e. , the word whose digits are ONE in any digit position where either or both of C(A) and C(X) contain a ONE. The application of an order of this type occurs principally in such logical processes as masking, in which only selected digits of a word are to be used in an operation. A more suitable order for this purpose is that for forming the logical product of C(X) with C(A); however, the operation LA X is so simply achieved that it is probably preferable unless a very rapid masking order is needed.

It is noted that LA X writes in the cores of Ad. If the order just preceding LA X were a clear and subtract operation, the cores of Ad would all contain ZERO's and the cores of Ac would contain the real "contents of A". It must be recognized that the pulse CL A places on the α sense lines a ONE wherever either Ad or Ac contains a ONE; i. e. , the logical sum of C(Ad) and C(Ac). Thus the effect would be the same as if it were Ad rather than Ac that contained the non-zero information. For all logical purposes, A may be regarded as a single row of cores in which a logical sum is formed as the writing takes place.

TABLE 2-4

LA X, Logical Add X, C(X) U C(A) into A, 8 Pulses

1 α CL S, CL D

2 to 8: The same as in Table 2-1

An operation having frequent convenience as a storage - saving device in arithmetic, but which in addition has an unique logical function in the present computer, is that of exchanging

C(X) with C(A). The logical function has to do with counter registers that receive input pulses from an external source. All counting (which uses the Z register) occurs at a definite place in the sequence generator shift register at which the Z register is vacant; specifically, this is between lines 4 and 5 in the previous tables. If the contents of an externally pulsed counter register are to be examined without risk of missing a pulse in the count, it is necessary to be able to pick up the counter contents and to reset the counter to a known value simultaneously in a single order. Even the order CA X, when applied to a counter register, is not a safe way of examining the register as the order is now written, since the pulses CL E and WE are separated by the time at which counters are incremented. The order EX X, Exchange with X, described in Table 2-5 does achieve the desired result, permitting the contents of the counter to be placed in A and those of A in the counter during a time period in which no incrementing can occur.

TABLE 2-5

EX X, Exchange with X, C(X) exchanged with C(A), 12 Pulses

- 1 α CL S, CL D
- 2 β (no action)
- 3 α CL E, WB, W1Sc
- 4 β CL B, WD
- 5 α CL A, WB, WP
- 6 β CL B, CL P, WE
- 7 α CL D, WB
- 8 β CL Z, WS, WD
- 9 α CL S, CL D, WZ, W1R2c
- 10 β CL B, WAd, W1Bc
- 11 α CL E, WB, WP, WR2c, W1Sc
- 12 β CL B, WE, WS, WD, CL Bc, CL P, TP

No parity check is made on the original contents of X in the exchange order; for counters, in the present mode of operation, a parity check is not kept. An alternate counter operation method in which the Z register would be on the α side of storage would permit parity to be preserved.

The exchange order is the first one that has been encountered which involves other than eight pulse times for its execution. Its beginning is identical to that for LA X and its ending is identical to all orders considered. In the interest of economy of cores and windings in the sequence generator core matrix, the shorter orders are therefore stretched out to match the twelve columns of cores required for EX X. This operation is not done in a way so as to cause the other orders to waste four pulse times of inactivity, but requires the shorter orders to skip over several positions of the shift register. The result permits each instruction to proceed with no waste of time, and yet permits many cases of sharing of matrix cores between different orders. The numbering of the lines in the tables shown here is for convenience only, and does not correspond to core matrix columns.

The instruction "Edit by X" is perhaps unique to this particular computer, and is motivated by the fact that many of the operations such as shifting are handled by special write wiring in registers in E storage. The function of this order, in brief, is to accomplish in one order of twelve pulse times what would otherwise require the two orders TS X, CA X and sixteen pulse times. Its operation is described in the table which follows.

TABLE 2-6

ED X, Edit by X, C(A) into X, then C(X) into A, 12 Pulses

- 1 α CL S, CL D, WB
- 2 β (no action)
- 3 α CL E, WlSc
- 4 β CL B, WS
- 5 α CL S, CL A, WB, WP
- 6 β CL B, CL P, WE
- 7 α CL E, WB, WP, WlSc
- 8 β CL Z, WS, WD
- 9 α CL S, CL D, WZ, WlR2c
- 10 β CL B, CL P, WAd, WE, WlBc
- 11 α CL E, WB, WP, WR2c, WlSc
- 12 β CL B, WE, WS, WD, CL Bc, CL P, TP

Orders dealt with up to this point have been chiefly concerned with moving information around from one point to another. Consider now the order TC X, transfer control to X, which causes the next instruction executed to be that in register X. In the process, the next consecutive address (i. e. , the address one greater than that of the location of TC X) is left in central register Q from which it may be stored by TA X, below. This pair of orders provides convenient entry and exit procedures for closed subroutines.

TABLE 2-7

TC X, Transfer Control to X, N + 1 into Q.

Next Operation from X, 6 Pulses

- 1 α CL Q
- 2 β CL Z, WQd
- 3 α CL S, CL D, WZ, WlR2c
- 4 β WlBc
- 5 α CL E, WB, WP, WR2c, WlSc
- 6 β CL B, WE, WS, WD, CL Bc, CL P, TP

It is noted that the order itself goes into the Z register in line 3, rather than merely the address, and a similar statement is true for certain branch orders considered later. This can, in turn, result in order codes being attached to the addresses which enter Q and thence are stored by the TA order. The effect of this, if recognized in programming, can cause little harm.

The transfer address order, TA X, serves an additional purpose beyond that of storing the contents of Q. This other purpose is the transfer to Q of the contents of R2c; thus two TA orders following each other will store successively the original contents of Q and of R2c. Otherwise, TA X closely resembles TS X in operation.

TABLE 2-8

TA X, Transfer Address to X, C(Q) into X, then
C(R2c) into Q, 8 Pulses

- | | | |
|---|----------|-----------------------------------|
| 1 | α | CL S, CL Q, WB, WP |
| 2 | β | CL R2c, WQ |
| 3 | α | CL E, CL D, W1Sc |
| 4 | β | CL Z, WS, WD |
| 5 | α | CL S, CL D, WZ, W1R2c |
| 6 | β | CL B, CL P, WE, W1Bc |
| 7 | α | CL E, WB, WP, WR2c, W1Sc |
| 8 | β | CL B, WE, WS, WD, CL Bc, CL P, TP |

We now come to consider a group of three branching orders, that cause a transfer of control to occur or not, depending upon the outcome of a test. Representative of such orders is BNZ X, branch non-zero to X, which tests central register A and acts as a TC X order if A contains any ONE's. Electrical details of the branching action have been considered only in a preliminary way; however, the principles are relatively simple. The original setting of a row of cores in the sequence

generator core matrix for BNZ X skips certain columns (i. e. , shift register positions). These are filled in subsequently in one of two alternate ways selected by the result of the testing operation. Thus lines 3 to 7 in Table 2-9 are filled with two choices which depend upon whether C(A) are zero or not. Control pulses describing details of this action are not shown; however, the testing is done as C(A) are written into B in line 1, and the actual action of setting one of two rows of cores is indicated by the word "BRANCH" in line 2.

TABLE 2-9

<u>BNZ X, Branch Non-Zero to X, Acts as TC X if C(A)</u> <u>are Non-Zero, Otherwise Order is</u> <u>Ignored, 10 Pulses</u>	
1 α CL S, CL A, WB	
2 β CL B, WA, BRANCH	
Active branch (non-zero)	Inactive branch (zero)
3 α (no action)	3 α CL E, WB, W1Sc
4 β (no action)	4 β CL B, WE
5 α CL Q	5 α CL D
6 β CL Z, WQ	6 β CL Z, WS, WD
7 α CL D, WZ, W1R2c	7 α CL S, CL D, WZ, W1R2c
8 β W1Bc	
9 α CL E, WB, WP, WR2c, W1Sc	
10 β CL B, WE, WS, WD, CL Bc, CL P, TP	

A branch order closely related to BNZ X is the order BMN X, branch minus to X, which tests the sign (digit 23) of A and acts as TC X if this bit contains a ONE denoting a negative number. The formal table of control pulses for BMN X is identical to that of Table 2-9 for BNZ X. The sole difference lies in one control pulse, not shown in the table, which deter-

mines the nature of the test. For BNZ X, all 23 bits of C(A) are tested, whereas for BMN X only the single high-order bit is involved.

A further branch order, similar to BNZ X except that the control pulses CL A, WB, CL B, WA are omitted in lines 1 and 2, is that for testing for overflow of a program counter. Certain erasable storage registers may be set aside as counters by provision of extra circuits described in Section E of Chapter 3. Those used by the program for its own bookkeeping purposes differ from those for external pulse inputs in the source of the input to be counted and in the disposition of any overflow pulses. Extra digit positions in the instruction word are so wired that a ONE in that digit causes an associated program counter to increment by one after the order is executed. For this particular class of counters, any overflow pulses are used to set an overflow core which may be tested at any later time. The "branch no overflow" instruction, BNO X, is the means for this test. This instruction acts as TC X provided that no overflow pulse has occurred, and is otherwise ignored (except that it clears the overflow core).

One final branching order of frequent convenience is that designated as TRSP X, "test register X and skip (next order) if plus". If C(X) are negative, the next order is taken from $N + 1$; if C(X) are positive, the next order is taken from $N + 2$. Further, the absolute value of C(X) is left in Q from which it may be stored by TA.

TABLE 2-10

TRSP X, Test Register X and Skip on Plus, 10 PulsesIf C(X) plus, next order from N + 2 and C(X)into Q; If C(X) negative, next order fromN + 1 and C(X)' into Q

	1	CL S, CL Q, W1R1c		
	2	CL Z, WD		
	3	CL E, WB, WP, WR1c, W1Sc		
	4	CL B, CL P, TP, WE, BRANCH		
Active branch (negative)			Inactive branch (positive)	
5	CL D, WB		5	CL D, WZ
6	CL B, WS, WD		6	CL Z, WS, WD, W1Qc
7	CL S, CL D, WZ, W1R2c		7	CL S, CL D, WZ, W1R2c
8	CL R1c, WQd, W1Bc		8	CL R1c, WQc, W1Bc
	9	CL E, WB, WP, WR2c, W1Sc		
	10	CL B, WE, WS, WD, CL Bc, CL P, TP		

The order CANP X stands for "clear and add, with no parity (check)". Its purpose is to provide a way to operate with registers for which the parity bit may be incorrect or absent; e. g., the register for shifting right one digit, described in Chap. 1, Section C, or an input register. The pulse sequence for CANP X is essentially the same as that of CA X in Table 2-1 except that the pulse WP is missing from line 3 and the pulses CL P, TP are absent from line 6.

The automatic interrupt, discussed in Chap. 3, Section C, is an interruption to the normal sequence of computer operations caused by various input circuits which require attention. It occurs, when required, just before execution of the last line in an instruction so that the next order has been called from storage but not yet transferred to the sequence generator. Instead of the next order, a special "interrupt" sequence of control

pulses is generated which, in addition to several specialized logical actions, stores away the contents of selected central registers required to permit the normal instruction sequence to resume later without any apparent discontinuity. These contents are stored in other central registers, not shown in Fig. 2-1; probably two such registers can be made to suffice.

The interrupt sequence is not itself a programmable order, however there is at present in the order code system an order RSM X, "resume" (X is irrelevant), which provides a control pulse sequence for resuming operation after an interrupt. The effect of RSM X is, first, to test whether another interrupt request has occurred while the actions called for by the original interrupt were in progress (no more than one interrupt is processed at a time). If not, the contents of the central registers are restored and the operation of the main program is resumed.

As any programmer knows, one of the most important computer capabilities required for compact and elegant programming is that of address modification; i. e., the capability of writing an instruction whose effective address at the time of execution depends upon other actions of the program itself. In this computer the problem is solved in terms of quite simple circuits by providing the capability for so-called "indirect addressing" for every order. The effect of indirect addressing on an instruction is to replace the address mentioned in the instruction by the address contained in that location. For example, let an asterisk following an order code indicate that indirect addressing is to be applied, and let the contents of register X be the number Y. Then CA* X is essentially synonymous with CA Y. If X is an erasable register (otherwise the operation is pointless) then its contents Y may be given different values by operation of the program, and the single wired-in instruction CA* X may thereby have a corresponding variety

of effective addresses.

When combined with the use of program counters and the previously mentioned capability of incrementing these automatically under control of selected digits in the order code, the indirect addressing capability permits some quite compact programs. For example, let it be desired to move N consecutive words in storage locations K to $K + N - 1$ to the new locations L to $L + N - 1$. To do this we place the address K in a counter $C1$, address L in a counter $C2$, and place the number $(N - 1)$ in a counter $C3$. Following this preliminary operation, the actual execution is accomplished with a loop of three orders:

► Address	m: CA* C1	and increment C3
	m + 1: TS* C2	and increment C1
	m + 2: BNO m	and increment C2

The electrical details of the indirect address feature are rather simple. Of the six bits specified for the order code, four are used for the fourteen orders thus far discussed. One more is reserved for the "extended operations" to be discussed shortly. The remaining one is for indirect addressed orders. Its effect, in brief, is to precede the control pulse sequence normal to the order in question by another of four pulses in length which replaces the contents of D and S by the word obtained from storage in the address specified by the order. The control pulse sequence for this purpose is given in Table 2-11. It is necessary in addition to cause the control sequence shift register to start at a position four steps preceding the beginning point for normal orders. This is accomplished through a relatively simple wiring in the sequence generator core rope. The whole cost of providing this facility is perhaps seven or eight transistors and a few cores, etc.

TABLE 2-11

Indirect Addressing

- 1 CL S, CL D
- 2 (no action)
- 3 CL E, WB, WlSc
- 4 CL B, WE, WS, WD

Normal operation of order follows

The so-called "extended operations" form an inexpensive storage-saving device by providing means for a one-word entry into a number of subroutines. As an example, multiplication exists in most variations of the computer only as a subroutine. To multiply $C(X)$ by $C(Y)$, the following orders might be written:

CA X
 TS to standard location
 CA Y
 TC to multiply routine

The extended operation version of this would replace the last three orders by one order. The problem is that four items of information are required: (1) address X or its contents, (2) address Y or its contents, (3) the address to go to after multiplication is complete, and (4) the fact that multiplication is desired; i. e., an address associated with this particular subroutine. A single order does not have sufficient digit positions to carry all of these data; two, however, can do so by relying on the Z and Q registers to handle the next address, provided that only a few bits are required for item 4.

The action, in brief, is that one bit in the order code causes a special control pulse sequence to become active, which holds $C(Z)$ in Q as in the TC order, preserves the complement of the original order in R2c, and transfers control to a specific address (e. g., address 128) which is generated electrically. There, a short standard program is initiated which files these

data away, including the original order, and finally transfers control to an address derived from the order code portion of the original instruction. The whole operation requires approximately 50 pulse times whereas the three-order version cited previously requires about half of that. Thirty-two of these extended operations are available. By and large they will be used for a variety of instructions found in most computers which are accomplished in ours by means of subroutines. There is some appeal, though, in planning a computer for geometric computations required, say, in space navigation, which may have "vector add" as a computer instruction.

D. Addition and Multiplication

The selection of an order code is dependent on the particular use to which the computer is put. The choices are, as usual, between speed and sophistication on the one hand and simplicity and reduced equipment on the other. Of particular interest in the present case is the method for addition, in which no use is made of an adder. Instead, in the representative computer, addition is carried out by a repetitive logical process within a loop in the sequence generator, in which all operations are conducted by the processes and registers already discussed.

The action is as follows. Let $u \oplus v$ stand for the bit by bit logical sum of two words u and v ; i. e., the word that contains a ONE in each digit position for which either u or v (or both) contains a ONE. To add two numbers x and y , the operations

$$a = x \oplus y$$

$$b = x' \oplus y'$$

$$C = b'$$

$$S = (a' \oplus b')'$$

are performed. The word S is the sum and C the carry for a one-digit-at-a-time summation of the words x and y . If C is zero, the process is complete. Otherwise, C is cycled leftward

one bit, with end-around carry, and S and the cycled carry word are used to replace x and y in the above formulae and the addition process repeated. The operation continues repeatedly until the carry word is zero, at which time the corresponding sum S is indeed the algebraic sum of x and y. The number representation is that of ONE's complements, in which $-x$ is the bit by bit complement of x.

The pulse by pulse operation of the loop portion of the control sequence is given in Table 2-12, in which the contents

TABLE 2-12

<u>Addition Loop</u>							
	Ac	Ad	Qd	Qc	CYc	R1c	R2c
	x	y	x'	y'	0	1	1
1 α CL A, WR2c, W1CYc, WB, BRANCH	0	0	x'	y'	1	1	a'
2 β W1Ac, CL B	1	0	x'	y'	1	1	a'
3 α CL Q, WCYc, WR1c	1	0	0	0	C	b'	a'
4 β CL R1c, CL R2c, WAc, WQd, W1Qc	S	0	S'	1	C	0	0
5 α W1R1c, W1R2c	S	0	S'	1	C	1	1
6 β CL CYc, WAd, WQc	S	C	S'	C'	0	1	1

of the relevant central registers are shown at successive stages of the loop. Starting with the two numbers to be added, x and y, in Ac and Ad before line 1, and their respective complements in Qd and Qc, in six steps these numbers are replaced by S, C, and their complements, and everything is ready for the cycle to repeat. The BRANCH operation in line 1 corresponds to a test

set up at line 6 of the previous cycle, in which Ad is wired to test for a non-zero carry word as C is written into it. The contents of A are written into B in line 1, so that when the **BRANCH** operation causes the control to leave the loop, the sum S is found in B.

The time required for execution of an addition by this process is approximately $14 + 6n$ pulse times, where n is the number of times S and C must be formed. A digital simulation of this addition process, using computer-generated "random numbers", has produced the plot of average n versus word length shown in Fig. 2-5. For each word length from 3 to 36, at intervals of 3, 4000 samples were taken to produce these data. Results show that for a word length of 23 digits, corresponding to the representative computer (neglecting the parity bit, of course), the average n is very close to five cycles. Thus approximately 44 pulse times, or 264 microseconds at 6 microseconds per pulse, are required on the average. The corresponding maximum time requires 24 cycles, giving 158 pulse times or 948 microseconds. Some care should be taken in making estimates based on the "average" value. The addition of a small negative number and a slightly larger positive number requires essentially the maximum time, since the carry must propagate all the way along the word one digit at a time.

We note in passing that the addition capability has been obtained quite inexpensively, given that the central registers shown in Fig. 2-4 already exist. Essentially, the only extra circuits required are those required for the branching operations, which involve perhaps eight or ten transistors altogether.

Roughly speaking, there are three ways of doing additions. The fastest and most expensive way in terms of equipment is to add with a parallel adder. With this scheme, addition of y to some x stored in A would require a single sequence, with no loops inside the sequence generator, of about 16 pulse times.

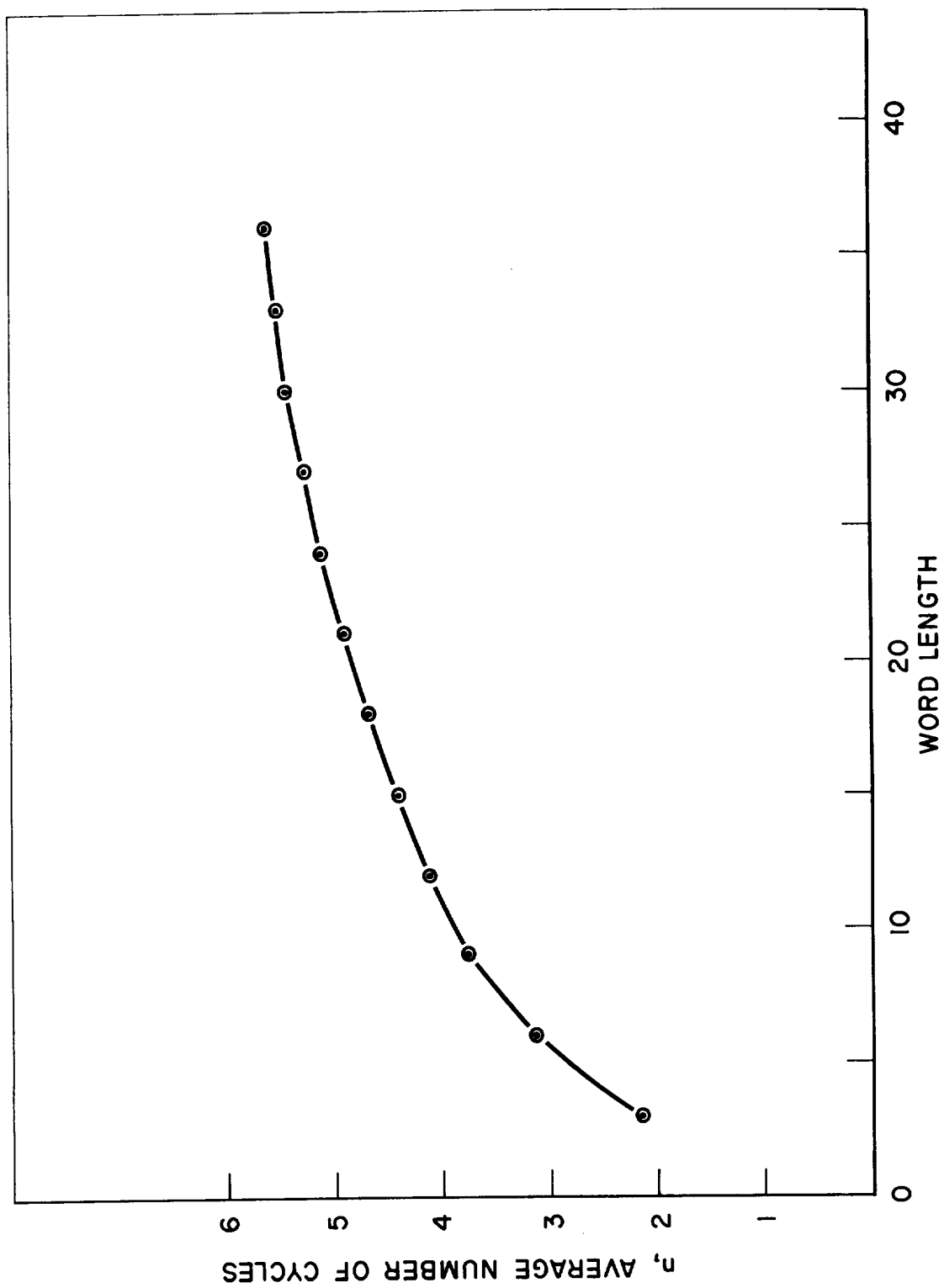


Fig. 2-5 Average number of cycles for addition.

The next fastest way to add was described in the above paragraphs, with a loop which is internal to the sequence generator. This method requires about 44 pulse times on the average, and 158 pulse times maximum. The slowest and most economical method is to perform addition with a program subroutine, using orders of which each is a zero level sequence with no loops inside the sequence generator, and with the orders stored in the core rope. This method is many times slower than the method described above, because in effect each pulse time of the microprogram sequence is replaced by an order which is itself some eight to twelve pulse times long. The saving in equipment consists mostly in the elimination of control pulse circuitry. There is little saving of registers because almost the same number of erasable storage registers must be reserved for addition by the program subroutine method as are used by the microprogram sequence. The difference is that in the subroutine method the registers are addressed by means of the S register, while in the faster microprogram method the registers are addressed directly by means of control pulses.

Multiplication in this computer is almost certain to wind up as an extended operation or subroutine. It should be remarked, however, that a "multiply loop" within the sequence generator is not out of the question, for an application in which the extra speed is necessary. No attempt has been made to draw up the logic for such a loop; however the following time estimates are probably reasonable for its operation. Use of such a loop surrounding a "one-shot" parallel adder circuit would perhaps require 200 pulse times for an average multiplication. If used in conjunction with an inner addition loop of the form discussed above, an average might be 500 pulses. These estimates may be compared to a value of from 1400 to 1500 pulse times for an average multiplication by subroutine methods, using the type of sequence generator addition loop discussed above. If addition, also, is a programmed operation, then of course the multiply order re-

quires much longer. It is thus seen that there are a variety of options open in the selection of a multiplication method. It is expected that most of the presently contemplated applications of this computer will be adequately served by one of the simpler procedures.

One rather unexpected consequence of the comparatively long time required for multiplication is the fact that, in those programs prepared to date, the operation time is easily estimated simply by counting the multiplications. Viewed in another light, the same proposition may be restated by saying that those operations of bookkeeping etc., which are incident to the preparation of an elaborate subroutine structure turn out to be quite inexpensive as far as computing time is concerned. As an example, the extended operation for vector addition requires about 400 pulse times; that for the dot product of two vectors requires about 5000. The amount of bookkeeping is identical in the two cases. The difference lies solely in the 4500 pulse times required for three multiplications. As a result, as long as this unbalance in operation time remains in this computer, it is quite likely that the compactness provided by use of numerous subroutine operations will more than compensate for the slight extra time this usually involves.

CHAPTER 3

INPUT AND OUTPUT SYSTEMS

A. Introduction

It is very difficult to discuss an input-output system without a clear definition of what the inputs and outputs are to be, and how many of them there are, and how fast they operate. As was mentioned in the introduction, it is difficult to describe a special purpose computer when the purpose, though special, is vague. In the present case, input-output systems will be discussed as a collection of techniques which the authors believe will be useful, in general, in the control of airborne and spaceborne systems.

As a background picture, it might be kept in mind that the origins of this computer were in a program studying guidance and control for an interplanetary probe. Representative problems for that application might include control of the vehicle orientation, through flywheels, in such a fashion as to align a body-fixed telescope tracker with the Sun or a star; or the positioning of a sextant to a precisely measured angle. Other tasks would include monitoring of various discrete signals describing the current states of various items of equipment. In general, various pulse trains are to be counted, representing the quantization of analog signals, and other steady state signals are to be examined. The rates involved never exceed a few hundred cycles per second, but the number and diversity of the data to be processed simultaneously may render an elementary scanning system difficult to achieve. The methods presented in

this chapter represent our approach to this type of problem.

To simplify discussion, it has been assumed that, as far as the computer is concerned, inputs from the outside world consist of on-off switches, mechanical or solid state. An output is similarly defined as a switch closure. The intent is to control such things as the angular position of a motor by on-off servo techniques. This sort of thing can be done if the computer response is fast enough. What follows is a set of schemes which will make the computer fast enough, in some cases at least, by logical means rather than by fast circuitry.

B. The Priority Circuit

The priority circuit (Fig. 3-1) is a novel device which permits certain economies in the handling of inputs. The circuit itself is related both to counters and to shift registers. In Fig. 3-1, let all cores be initially in state ONE. When transistor TA is saturated, core (1) starts to switch, and in so doing it causes transistor T1 to switch. This in turn causes all current to flow through T1, with almost no current flowing through the windings of cores (2), (3) and (4). It is assumed that the saturation resistance of T1 is zero compared to R2. It is further assumed that TA is turned off at the same time that core (1) finishes switching.

The next time an advancing pulse is applied to the base of TA, core (1) will be at zero; cores (2), (3) and (4) at ONE. This time core (2) switches in exactly the same fashion as did core (1).

If all the cores of the chain start at ONE, then the chain behaves much as a shift register with a single ONE travelling down it, in the sense that the numbers of cores in the chain determine the number of pulse times necessary to make the last core switch. In the case of the priority circuit, however, it is possible to reduce the effective length of the chain by not starting with all ONE's. For example, if core (3) had started at ZERO, then it would be the third advancing pulse instead of the fourth,

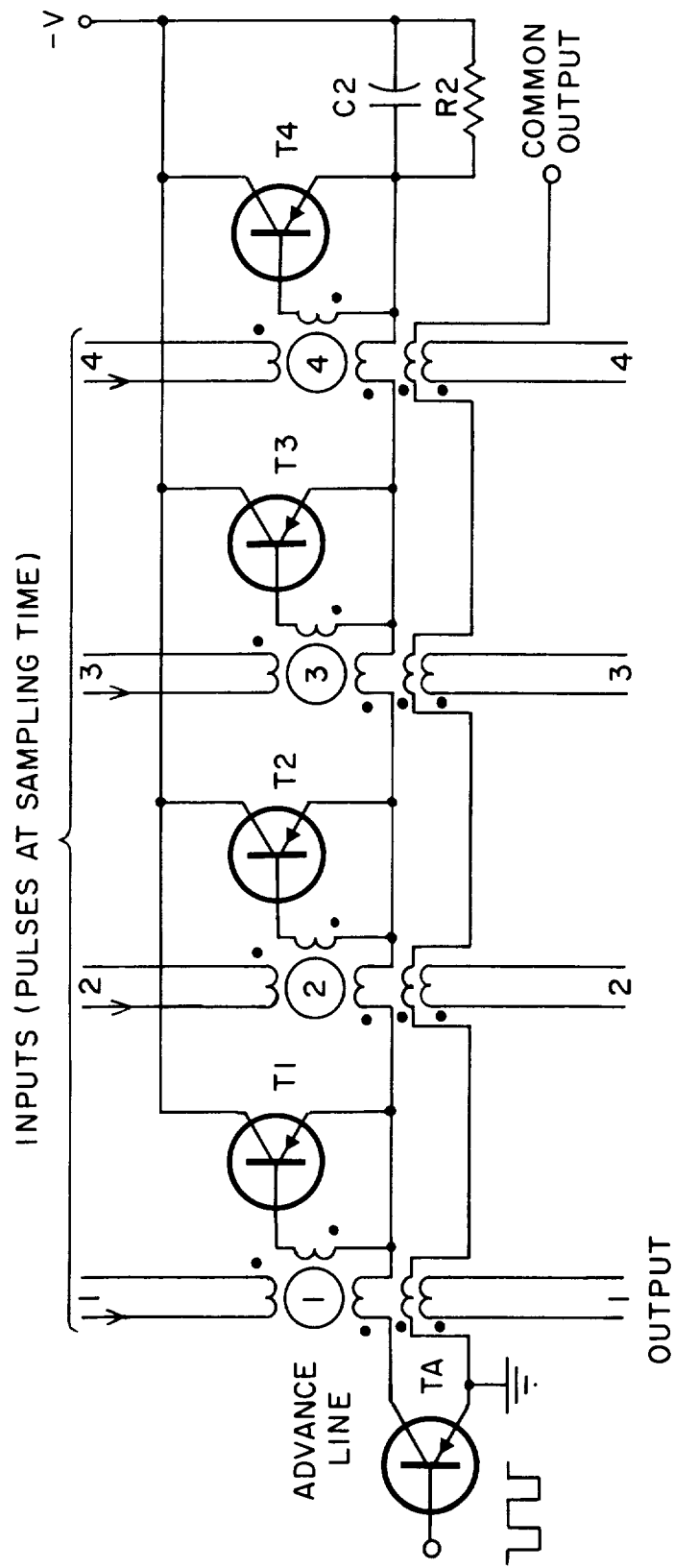


Fig. 3-1 Priority circuit.

which causes core (4) to switch.

Consider next a step of input leads such that a pulse on an input lead sets a core to a ONE (top of Fig. 3-1). Input pulses can only occur at a specific sampling time T_s . A winding common to all the cores in the priority circuit (labelled common output) can be used to sense whether or not one or more inputs occur during a sampling time T_s . If one or more inputs do occur, then the pulse which appears at the common output can be used, after a suitable delay, inversion, and reshaping, to trigger TA, Triggering TA results in driving from a ONE to a ZERO, the first core to be in state ONE, which in turn creates outputs at both an individual output winding, and at the common output. This last pulse, again delayed and reshaped, but this time not inverted, can be used to trigger TA once more. The process, schematized in Fig. 3-2, will continue until all the cores in the priority circuit are at ZERO, and will require as many advancing pulses as there were inputs at T_s (plus one). The name "priority" arises from the fact that input line 1 is always served (i. e., core (1) cleared to ZERO) first if it has an input, and input line 2 next, so that line 1 has priority over line 2, and so forth.

The sampling system of Fig. 3-2 is shown in more detail in Fig. 3-3. The inputs to the priority circuit are pulses generated by memory units M indicating that a switch has changed state since the last sampling time. This type of input, rather than a type which gives forth a pulse for every "on" switch every sampling time, permits an input system with some advantages over conventional input scanning systems. Furthermore, the memory units M of Fig. 3-3 can be so designed as to provide information on the actual state of the corresponding switch, and thus avoid the dangers of a pure "change of state" system, which is vulnerable to loss of pulses. A realization of unit M will be discussed later.

It is interesting to compare the priority circuit input

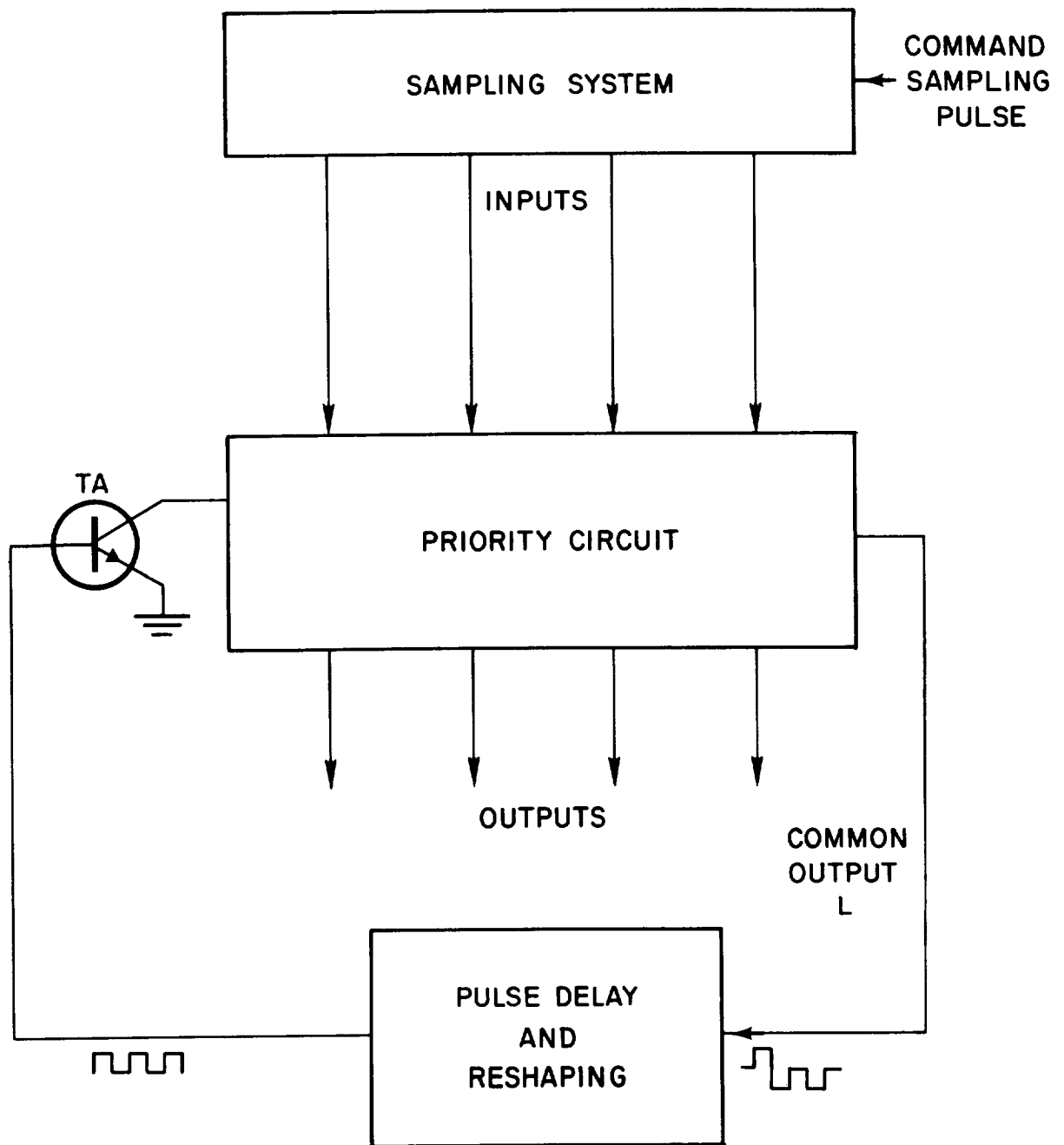


Fig. 3-2 Priority circuit system.

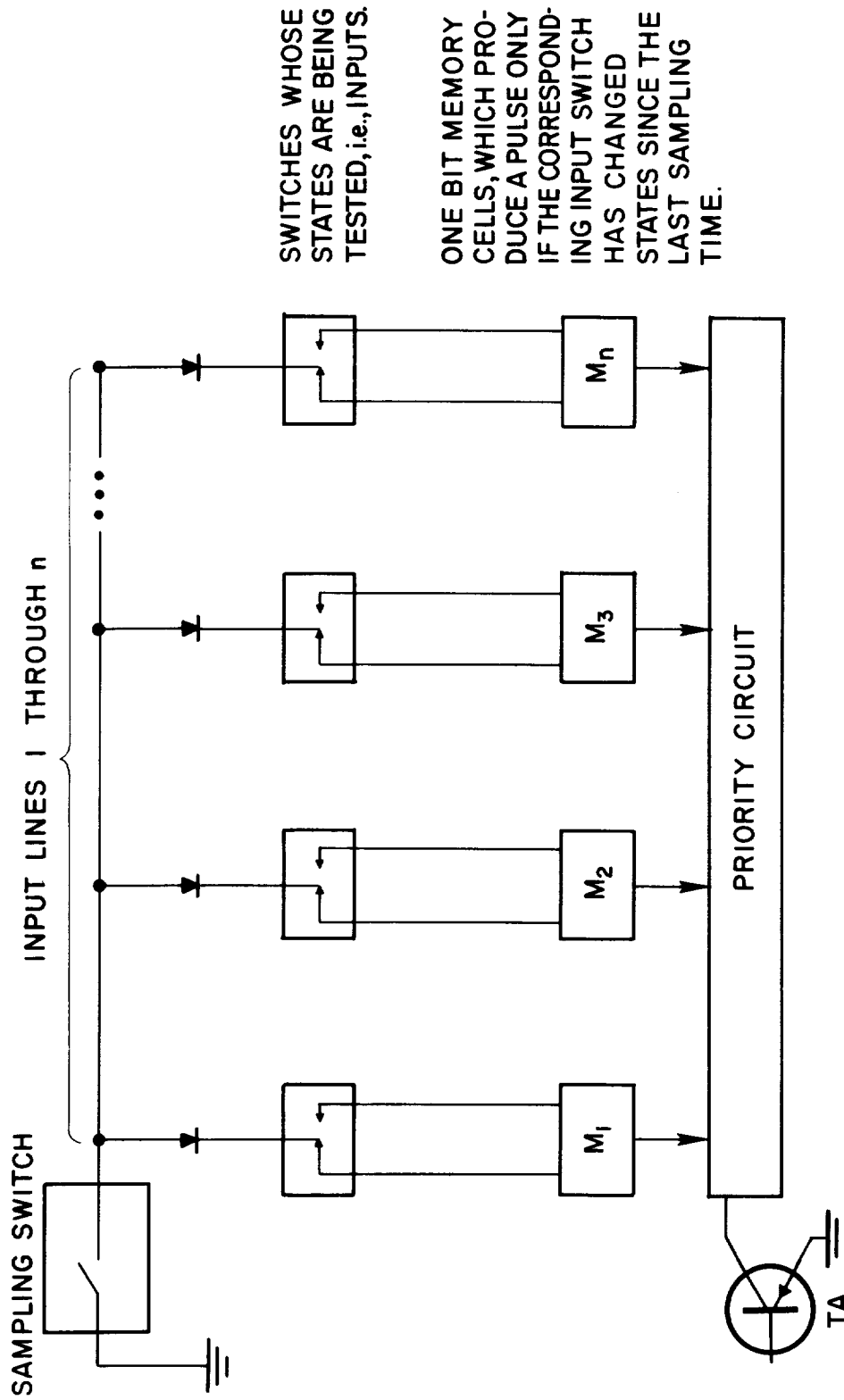


Fig. 3-3 Input circuits for priority circuit system.

system with one in which there is a shift register in place of the priority circuit proper. For the shift register, if there are N input lines, N advancing pulses by transistor TA are required between sampling lines. This second system in effect scans the N lines in between sampling times. If there are K sampling times per second, then the overall pulse handling capability of the second system is KN pulses per second. Assuming some technological limit to the pulses per second which may be handled by a given type of component, then the maximum sampling rate $1/K$ is determined essentially by the number of input lines to be served. Furthermore, each line separately may handle no more than K pulses per second* with this arrangement.

With the priority circuit in place of the shift register, it is no longer necessary that there be at least N advancing pulses on transistor TA in between every sampling time. The number of such pulses is determined by the overall activity of the input lines rather than number of such lines. Hence, advantage may be taken of any properties of the input lines, such as, for instance, that they be many in number but that their average activity be low.

More important perhaps is the ability to increase the sampling rate to keep pace with the fastest of the input channels, without necessarily requiring the advancing pulses on transistor TA to speed up proportionately. Suppose for the sake of definiteness that the processing of one pulse requires 200 microseconds and that, of 30 input channels, 6 may have up to 400 pulses per second whereas the rest may have no more than 10 pulses per second each. A maximum of 2640 pulses per second is therefore involved. Now if a scanning system requires 200 microseconds

* In speaking of the pulse carrying capacity of a channel in terms of pulses per second, we, of course, are really referring to the reciprocal of this number which is the minimum time interval between consecutive pulses.

to look at each channel, whether a pulse is present there or not, it cannot cope with the situation as described. To do so would require a processing rate of $30 \times 400 = 12,000$ pulses per second. Of course two separate scanning systems could be employed, one to scan the six high frequency channels and the other to handle the rest.

In a sense, the priority circuit input system acts in this way, except that it adapts itself automatically to varying input requirements. All that is required is to make the sampling rate fast enough to match the fastest of the input channels and, simultaneously, to be able to process pulses fast enough to handle the total load. Thus, in the example cited, a sampling rate of 500 per second would permit ten advance pulses on the priority circuit during each sample period (at 200 microseconds each), which would permit the six 400 cycle lines to be processed once each sample time and still leave time for handling four other pulses distributed among the remaining 24 lines. It could, of course, happen that more than four of the remaining 24 inputs would require processing in a particular sample period. However, the priority circuit would carry this information over until the next sample period, and could be guaranteed to process each channel before the next pulse on that channel arrived. From the standpoint of the tie-in to the computer, discussed in the next section, the priority circuit provides a means whereby the computer gives exactly the minimum possible amount of attention to its inputs, thus permitting highly efficient time sharing procedures.

A careful study of the properties of the priority circuit input system is beyond the scope of this report, but it seems clear that such a system can take advantage of variations in information rate among the input lines, and that it is in no case worse than the scanning system.

The memory unit M mentioned earlier may be realized

by the circuit shown in Fig. 3-4. Upon closure of the sampling switch, cores A1 and A2 are driven to ONE for one position of the input switch, and to ZERO for the other. Cores B1 and B2 are driven to ZERO and ONE, correspondingly. Cores A2 and B2 are always in opposite states, and will change states only if the input switch has changed states since the last closure of the sampling switch. A full wave rectifier generates a negative pulse every time A2 and B2 change states, and this pulse constitutes the input to the priority circuit.

It may be necessary to ascertain the state itself, rather than a change of states of the input switch, and cores A1 and B1 can be used for this purpose. These cores, which may be part of an erasable storage register, may be both cleared to state ZERO simultaneously, and their outputs sensed. Note that only one of A1 or B1 may be at ONE. The next closure of the sampling switch restores the proper core to a ONE, in correspondence with the state of the input switch. This subject is developed further in the next section.

C. The Automatic Interrupt

The automatic interrupt system provides integration of the priority circuit input system with the rest of the computer. The desired overall behavior of the input system is approximately as follows: when an input occurs a signal is given to the sequence generator and the normal chain of events is interrupted at the end of the current sequence. The next sequence (i. e., order) is not the order determined by the contents of register B, but a special sequence called the Automatic Interrupt Sequence (AI). This sequence transfers the contents of selected central registers for safekeeping into other central registers specifically reserved for this purpose. Then, by mechanisms to be described later, program control is transferred to a subroutine appropriate to the particular input which activated the interrupt. When such action is ended, a Resume sequence restores the contents of the central

registers and program control is transferred back to the instruction displaced by the AI sequence. In short, the occurrence of any input results in the automatic interruption of whatever the computer is doing, preservation of the contents of certain central registers, processing of the input, and then resumption of the previous computer activity. There is an exception, namely that it is not possible to interrupt a program which was entered into because of a previous interrupt. If an input B occurs while processing another input A, or at the same time as another input A with higher priority, then input B is processed after A is processed, and before the computer resumes its previous activity. A further exception lies in the fact that the interrupt action can be inhibited (i. e. , delayed) by the presence of a ONE in a designated position in the order code. Assuming this facility to be carefully used by the programmer, certain central registers (e. g. , Q and R2c) need not be preserved by the AI and Resume sequences. Thus some equipment savings result.

Roughly speaking, inputs may be divided into two categories: those which indicate the state of a switch, and those which indicate the presence or absence of a pulse, where the state of the switch which generates the pulse is not itself important. For example, the switch which indicates that a rocket motor is on is an input of the first kind. The train of pulses which comes from a flywheel pick-off is an input of the second kind. The point is that inputs of the second kind feed into counters, and that, in terms of input circuitry, cores A1 and B1 of Fig. 3-4 can be eliminated. The integration of each kind of input with the computer will be treated separately. The first kind will be called State Inputs, and the second kind Counter Inputs.

D. State Inputs

State inputs are handled by associating one (or more) input switches with an addressable register. This register is a combination of the erasable and the fixed kinds. Such a register,

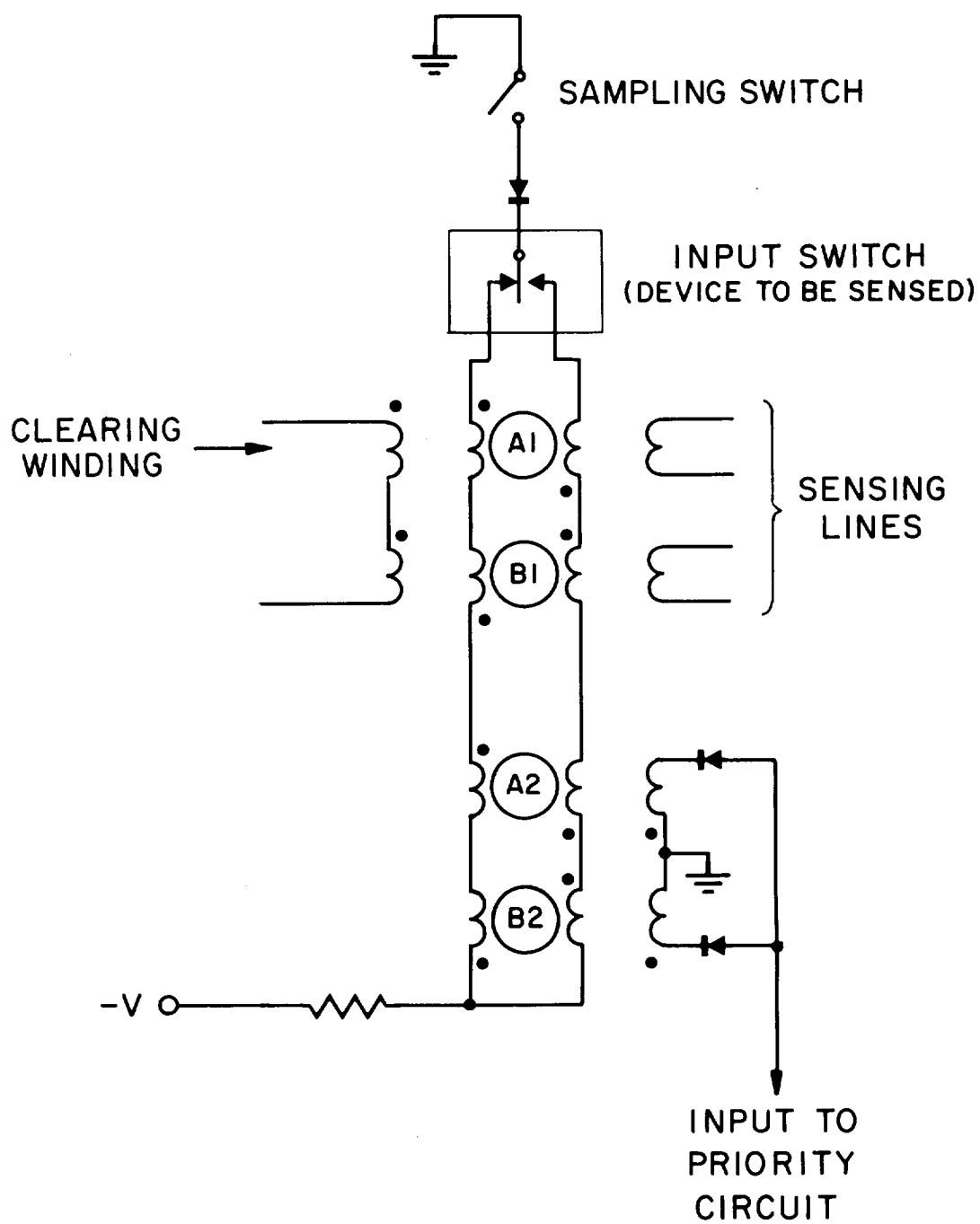


Fig. 3-4 Unit M for input to priority circuit.

with some associated circuitry, is shown in Fig. 3-5. The overall circuit is an elaboration of the circuits of Fig. 3-3 and 3-4. The addressable register itself is composed of cores R1, A1, B1, and R2. Core R1 is a rope core, and can be set to a ONE by having the proper address cleared from the S register, as usual, or by the occurrence of a change of state of cores A2 and B2, as will be explained presently. Cores A1 and B1 describe the state of the input switch, as in the circuit of Fig. 3-4. Core R2 follows the state of R1, and is threaded by some of several sensing lines. (This core could actually be eliminated, by threading the lines through R1 itself.) All of these sensing lines go into the A Sense Write amplifiers shared by E storage.

Operation is as follows: A change of state of the input switch from one sampling time to the next results, at sampling time, in a signal on the base of transistor T_P . This signal was previously described as "input to the priority circuit" (Fig. 3-4), and its net result is to drive to a ONE the priority circuit core P. The act of switching core P from ZERO to ONE also results in a signal on line L, which is common to all priority cores, and which goes to the sequence generator. The sequence generator then goes into the Automatic Interrupt sequence which generates, among other things, an Advance pulse. This advance pulse drives to ZERO the highest priority P core, say the one of Fig. 3-5. In so doing, transistor T_R generates a pulse which drives to a ONE the rope core R1, and the register core R2. When the sequence generator produces a CLE pulse during the resulting AI sequence, core R1 is driven to ZERO, and all cores of its associated register, A1, B1 and R2 in this case, are driven to a ZERO. In so doing a word is read into B which consists of the two bits which indicate the state of the switch, and other bits which give the address of the subroutine pertinent to the input being processed. The AI sequence then separates the word into two different words, and transfers program control to the input subroutine. When that subroutine is finished, a Resume sequence is generated, and another advance pulse is produced to test the

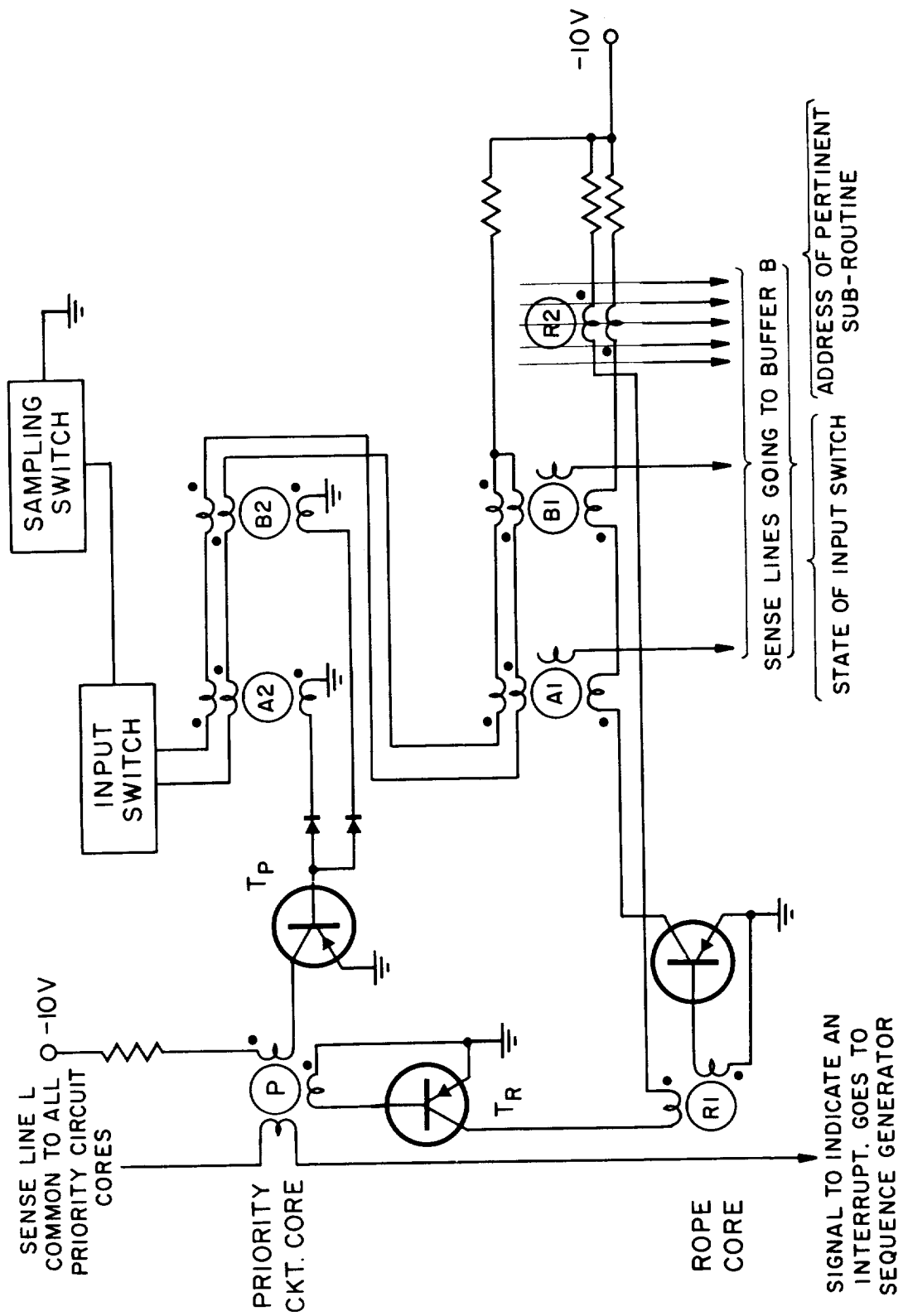


Fig. 3-5 Circuitry for state input.

presence of other P cores which might have been driven to a ONE. Assuming no other inputs, occur, the computer returns to its previous task.

Cores A1 and B1, cleared by a signal from the transistor associated with R1, are restored to correspond to the input switch at the next sampling time after action is initiated. In so doing, one of A1 and B1 changes state from ZERO to ONE, and for this reason they cannot be used to generate the "change of state" signal to transistor T_P. If they were so used, the CLE pulse, which clears the register R1, would generate a false interrupt signal, setting P to a ONE.

In the example given above, an entire register is used for a single input switch. Obviously, more than one switch may be associated with each register, requiring some additional circuitry for recording which switch caused the alarm. As a final remark on State Input, registers, it should be noted that these registers do not have the means for writing into them possessed by normal erasable storage registers. They can be written into only by means of input switches. The state of the input switch may be examined at any time, under program control, since register R1 can be addressed by normal means. Since no Write busses are provided for cores A1 and B1, however, this interrogation should not be made oftener than once per sample time.

E. Counter Inputs

Counting is an indispensable and frequent activity of control computers of the type under discussion. The sources of things to be counted are both internal and external. In the case of a space vehicle, the external sources are things such as accelerometer outputs, pulses generated by the rotation of a flywheel, a clock which keeps time. Internal sources are various bookkeeping operations necessary for a complex program, such as the number of iterations of a converging process,

or index modification in the case of programs dealing with matrices or vectors.

In either case, a solution which permits programming and computing time economies is to have special registers. These ideal registers have all the attributes of addressable erasable storage registers plus two others: they count pulses on an input lead independently of other computer activity, and they generate a signal upon overflow (Fig. 3-6). The contents of an ideal counter register may be examined at any time by reading said contents into B, and a word may be transferred from B into the counter; this last property allows counters to be preset to any number, under program control.

It is possible to realize these ideal counter registers or ones very nearly ideal by having active counters, with many components. A more desirable solution, where many counter registers are to be employed, is to have a system in which the registers themselves are passive, i. e., cannot do any counting, and, where there is a central unit, to add one to a number read into it which is shared by all counter registers. It will be recalled that such a unit was described in the section on central registers, i. e., the Z register. This solution is more desirable only because it saves equipment. The properties of the counter registers become slightly less than ideal, because of certain timing hazards which arise from the use of common equipment, and because the overall pulse handling capacity is limited to the capacity of the Z register. However, the end product is a reasonably economic system for a few active inputs, and a very economic system for a large number of slow inputs. As will be seen, there are many similarities between the proposed input counting system, and the more general state input system described previously.

Shown in Fig. 3-7 is a schematized system of two counter inputs and counter registers. When an input occurs, the priority

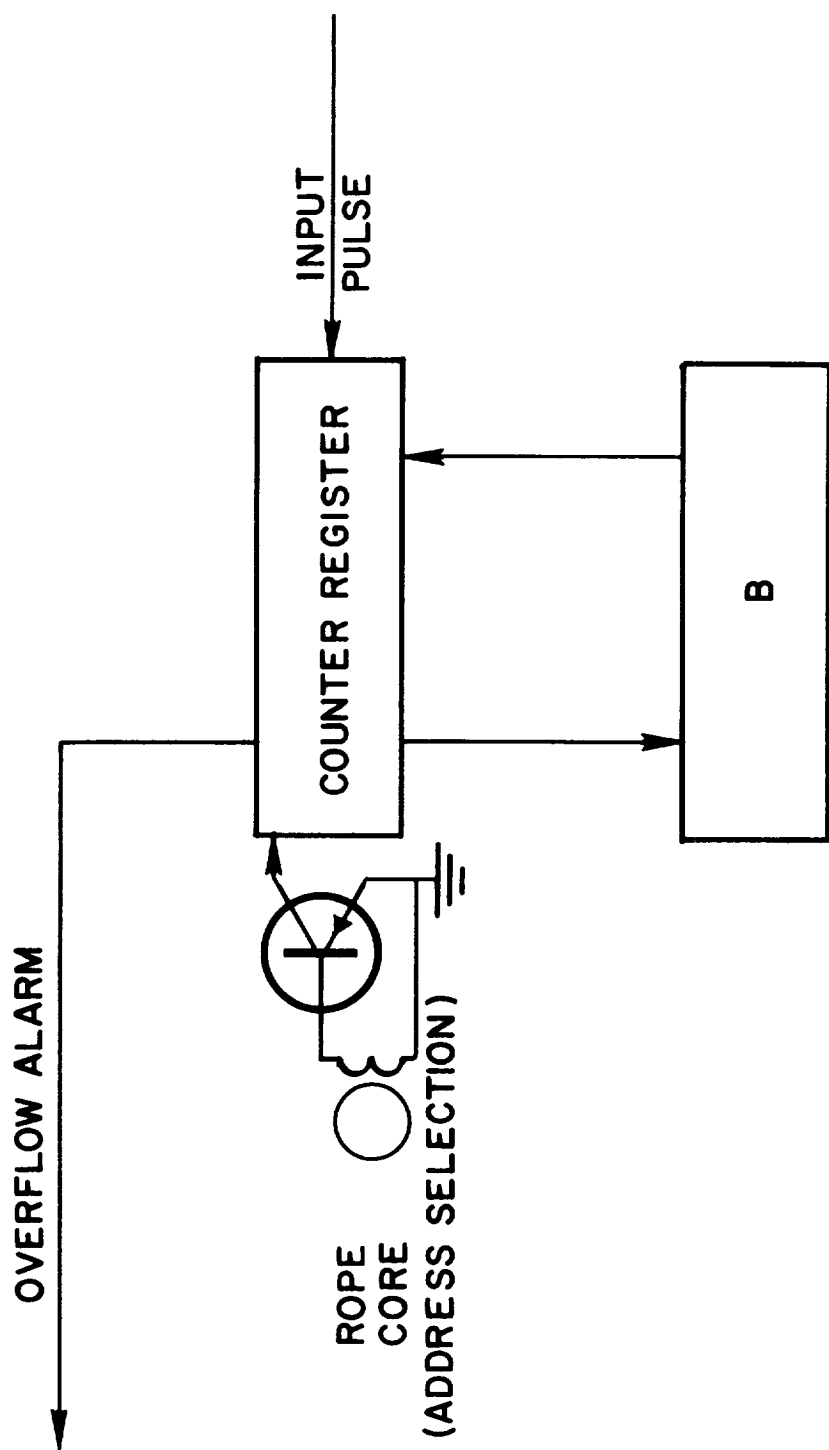


Fig. 3-6 Ideal counter register.

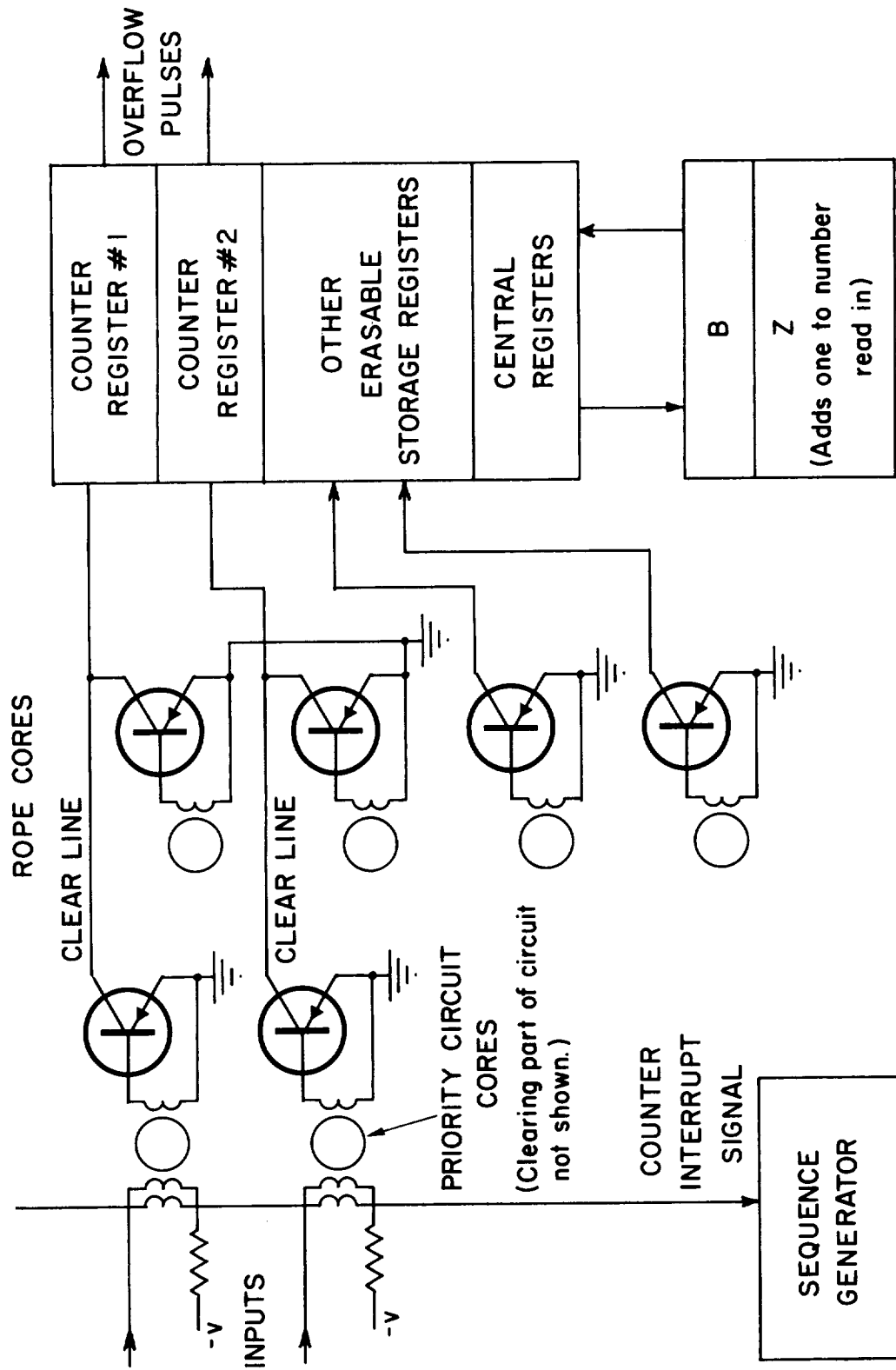


Fig. 3-7 System for counter inputs.

circuit core associated with that input is driven from ZERO to ONE, and an interrupt signal is generated. As in the case of state inputs, an interrupt signal causes the sequence generator to enter a special sequence. However, the special sequence which applies to counter inputs is shorter and simpler than that which applies to state inputs. Specifically, an Advance pulse is generated, which clears the priority circuit core to ZERO, and at the same time, a WZ pulse is generated. Clearing the priority core causes the Counter Register associated with it to be cleared to all ZERO's (just as if the proper rope core had been set to a ONE) followed by a CLE pulse. Because of the WZ pulse, the contents of the counter are read into Z, where a ONE is added in the least significant place.

The very next pulse time, the pulses generated are CLZ and WE, which cause the incremented count to be placed back in the proper Counter Register. Following this, another Advance pulse is generated to test for unserviced inputs, and the two pulse time process repeated until there are no more priority cores which have to do with counters at ONE. When this condition is met the sequence generator returns to normal functioning. All counters are updated at the end of every sequence, and once per cycle in the addition loop.

It is worth while to look closely at the difference between the processes for state and counter inputs. The most striking difference is in the times required to process each kind of input. A state input interrupts the program currently being executed, and transfers control to a special subroutine which deals with that input. This requires that many orders be executed, and since each order is about 10 pulse times long, each state input requires hundreds of pulse times. A counter input, on the other hand, requires exactly two pulse times to be processed. These pulse times are sandwiched between pulses of a normal sequence, when the Z register is free. Thus, the

current program is interrupted only in the sense of a short delay being introduced. Further, the circuits may be so arranged that no delay whatever is experienced when there is no counting to be done.

Each Counter Register has a transistor associated with its most significant digit core (not shown in Fig. 3-7), which generates an overflow pulse when the count reaches a certain size. This pulse can become a state input, so that a special subroutine is entered into every time that counter overflows. Alternately, the inputs to a counter can come from within the computer itself. For example, a rope core, whose switching causes the execution of an order, can have a transistor associated with it which feeds a counter. Every time that order is executed, the contents of the counter are automatically increased by one. This permits programs to be efficient in the sense of little time being wasted on some aspects of bookkeeping.

The counters described thus far can only be used for accumulating positive increments. There are basically two alternatives to the problem of positive and negative increments. One of these assigns two counters to each variable capable of increments of either sign, and a separate count is kept both positive and negative increments. This arrangement, while workable, has the undesirable feature of requiring a subtraction every time the absolute count of a variable is to be examined.

The other solution to the problem of bipolar increments associates two priority input cores with each counter, one for positive and one for negative increments. Then, depending on which input core is excited, either an "Add One" or a "Subtract One" sequence is entered. The "Add One" sequence can be as before. The "Subtract One" sequence can still be done with only the Z register by a double complementing operation, since complementing the contents x of a counter, adding one to the result, and complementing this in turn is the same as subtracting

one from x [$(x' + 1)' = x - 1$; $x = x'$]. This scheme has the disadvantage of requiring four pulse times to increment a counter, and also that the Z register be placed on the α side, instead of the β side where it is presently assigned (Fig. 3-8). The new location of Z in turn requires a different incrementing sequence for positive increments than the simple two pulse sequence used in the case of positive increments only. The two sequences are listed below.

Positive Increments

- 1 CL CNTR, WC1
- 2 CL C1, WZ
- 3 CL Z, WC1, WP
- 4 CL C1, CL P, WE (= W CNTR)

Negative Increments

- 1 CL CNTR, WC2c, W1C3c
- 2 CL C2c, WZ
- 3 CL Z, WC3c, W1C2c, WP
- 4 CL C3c, CL P, WE (= W CNTR)

Registers such as C2c are complementing registers. The above four pulse sequences have the advantage of permitting a parity bit in counters, which is updated automatically, provided that the number of bits in a word, excluding the parity bit, is even. The two pulse sequence, of course, disregarded parity completely. The condition of an even number of bits is to insure that the parity of a word and its complement be the same.

A final and most desirable solution to the problem of bipolar counting involves a modified Z register capable of subtracting one (by adding the complement of one), as well as adding one to the quantity entered into it. Such a modified Z register is currently under development, and it may be possible to build one so that it still requires only a single transistor per bit.

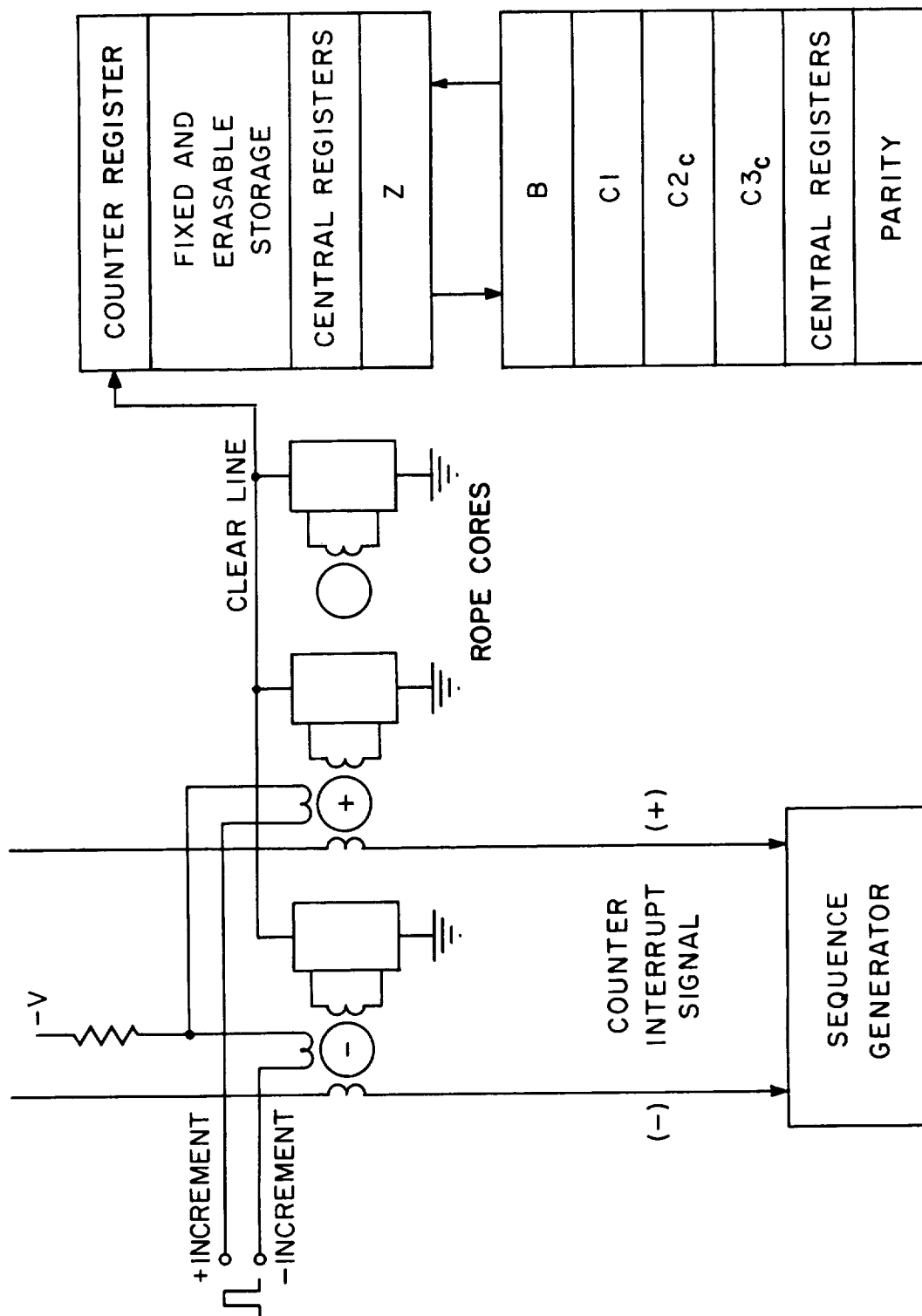


Fig. 3-8 Counter inputs--system for dealing with positive and negative increments.

F. Output Devices

As was the case with input devices, output devices will be discussed without any specific application, and presented as a useful collection of ideas, circuits, and subsystems. No part of the problem of digital to analog conversion will be discussed here, but the generation of power pulses (as opposed to pulses used for signalling purposes only) will be mentioned.

The backbone of the output system is the core rope. Cores in the rope may have transistors or a latch type of solid state device connected to them. The core may be addressed by means of the S register, just as any other fixed or erasable register. Switching one of those cores then generates a pulse, or causes the closure of a switch which remains closed until some other core switches.

The latch devices mentioned above are of particular interest because they approach the ideal relay contact; they are three terminal, four layer transistors which turn on when the base is pulsed positively relative to the emitter (Fig. 3-9), remain on as long as a sustaining current flows through them, and are turned off by pulsing the base negatively relative to the emitter. The sustaining current is of the order of 10 ma. The "on" and "off" signals are of the order of 1 to 5 volts, and about 1 μ sec in duration. As shown in Fig. 3-9, these latches can be used to keep a transistor turned on. When the latch is off, the base of the transistor is held at +5 volts. If core 1 fires, the latch turns on, and the base of the transistor is held negative relative to the emitter. The latch is turned off by core 2. Latches may be synthesized out of ordinary transistors, if need be.

An application of the output switch circuit is shown in Fig. 3-10, where it is used to gate a pulse generator. Thus the load may be subjected to trains of pulses, or rate R, under program control. The pulse generator may be common to several output switches.

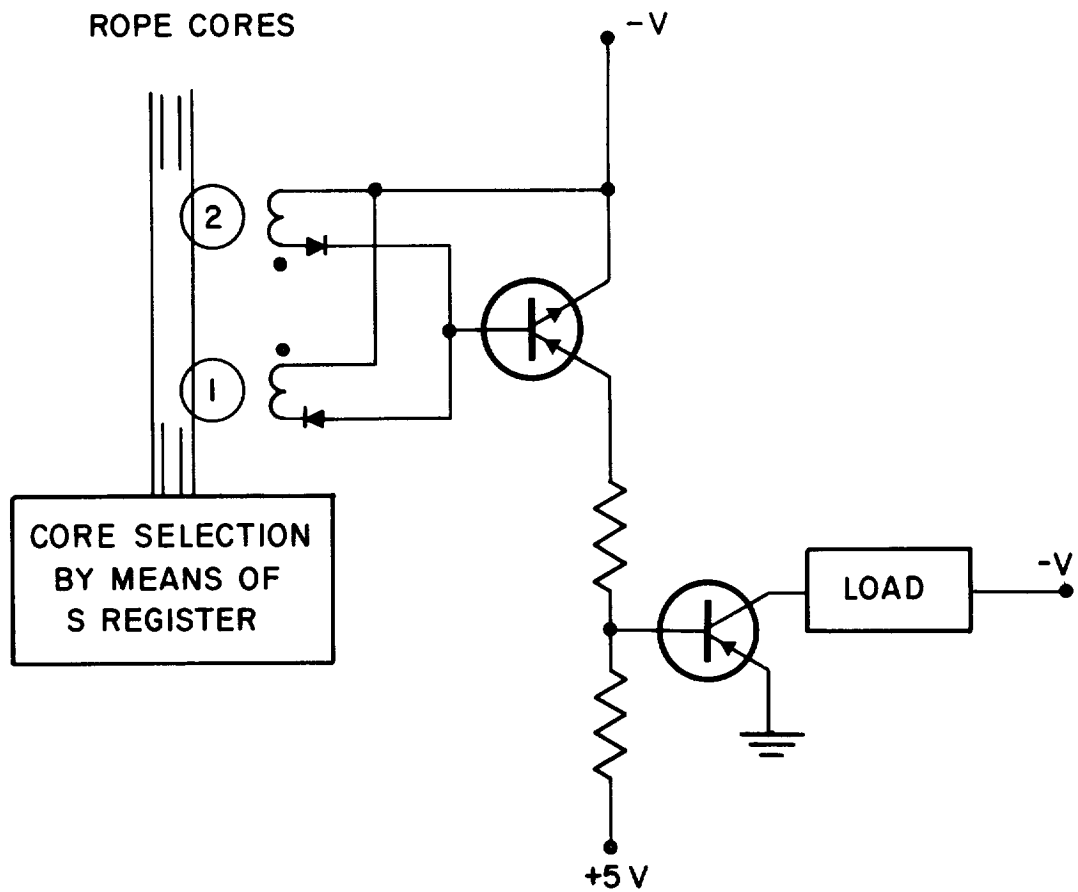


Fig. 3-9 Output switch circuit.

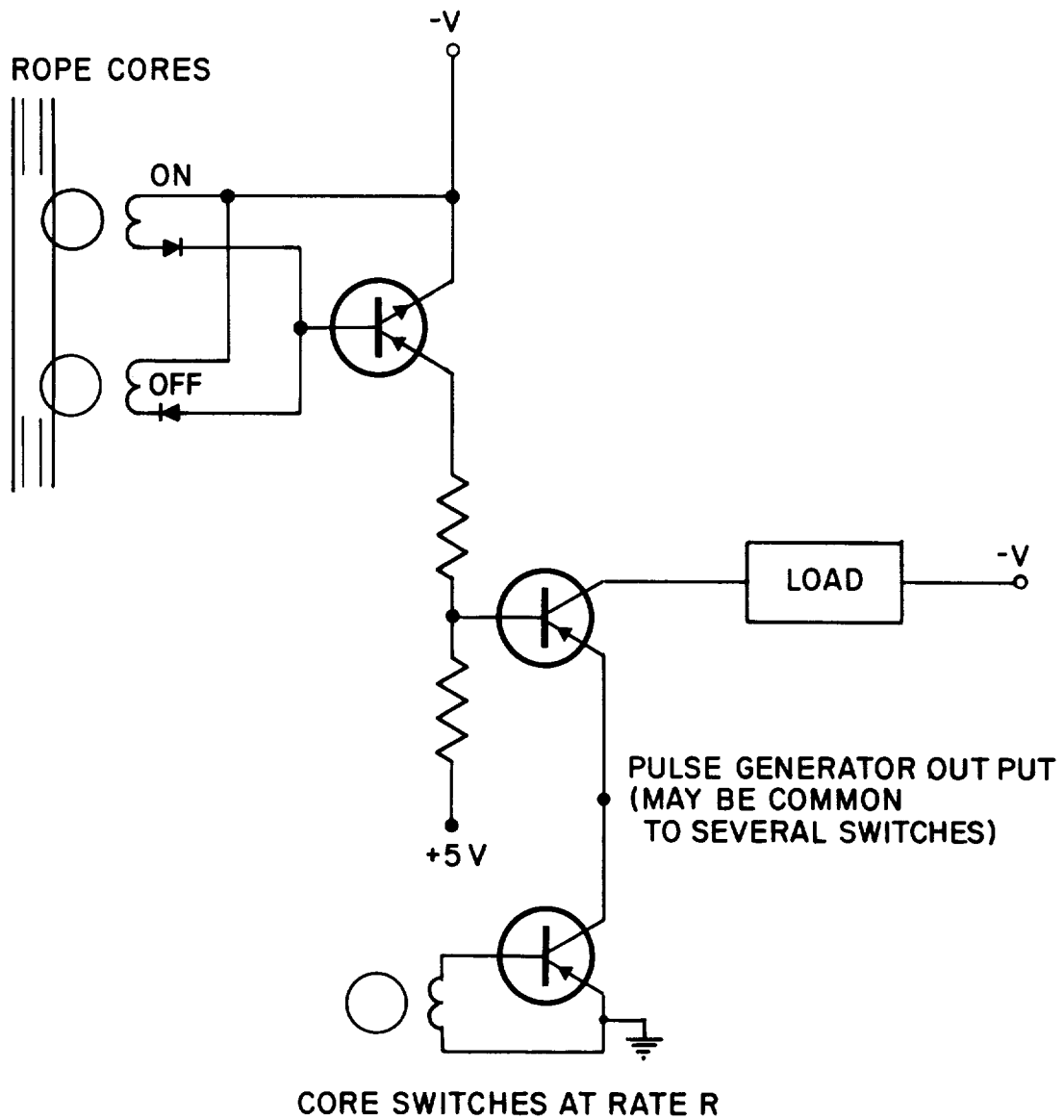


Fig. 3-10 Gating of a pulse generator by an output switch.

An important output problem is that of pulse rate generation. From the point of view of efficient computer utilization, it is desirable that there be a kind of pulse generator whose output rate is set by the computer by simply writing a number into a special register. The intent is to free the computer from the task of regulating pulse rates. The computer sets a rate in a given pulse generator, and this rate is produced independently of the computer until the rate is next set. The premise is that the intervals between settings of the pulse rate are large compared to the intervals between pulses. With this type of pulse generation it is possible to use a medium speed computer for the overall control of many independent pulse rates, each of which varies slowly, but may be of the order of the computer clock rate.

The desired pulse generator features may be achieved at relatively low cost. The basic scheme is outlined in Fig. 3-11. There is a counter associated with a clock (which could be the computer clock), out of which come the various pulse rates $R, \frac{R}{2}, \frac{R}{4}$, etc, each pulse rate appearing on a separate wire. For each independent pulse rate Q to be generated there is an Output Rate Register. These registers are part of the E storage, and their contents can be set by the computer. Each bit position of each register acts as a gate for one of the clock counter rates, as shown in Fig. 3-12. When a ONE is read into the core corresponding to a bit position, the solid state latch tied to that core is turned on, and remains on until the core is cleared to a ZERO. While the latch is on, the base of the transistor connected to the latch is held negative, and the transistor acts as a closed relay contact. While the latch is off, the transistor does not conduct because its base is held positive.

A desired rate Q is generated by combining pulse rates from the clock counter. Thus

$$Q = a_1 R + a_2 \frac{R}{2} + a_3 \frac{R}{4} + a_4 \frac{R}{8} + a_5 \frac{R}{16}$$

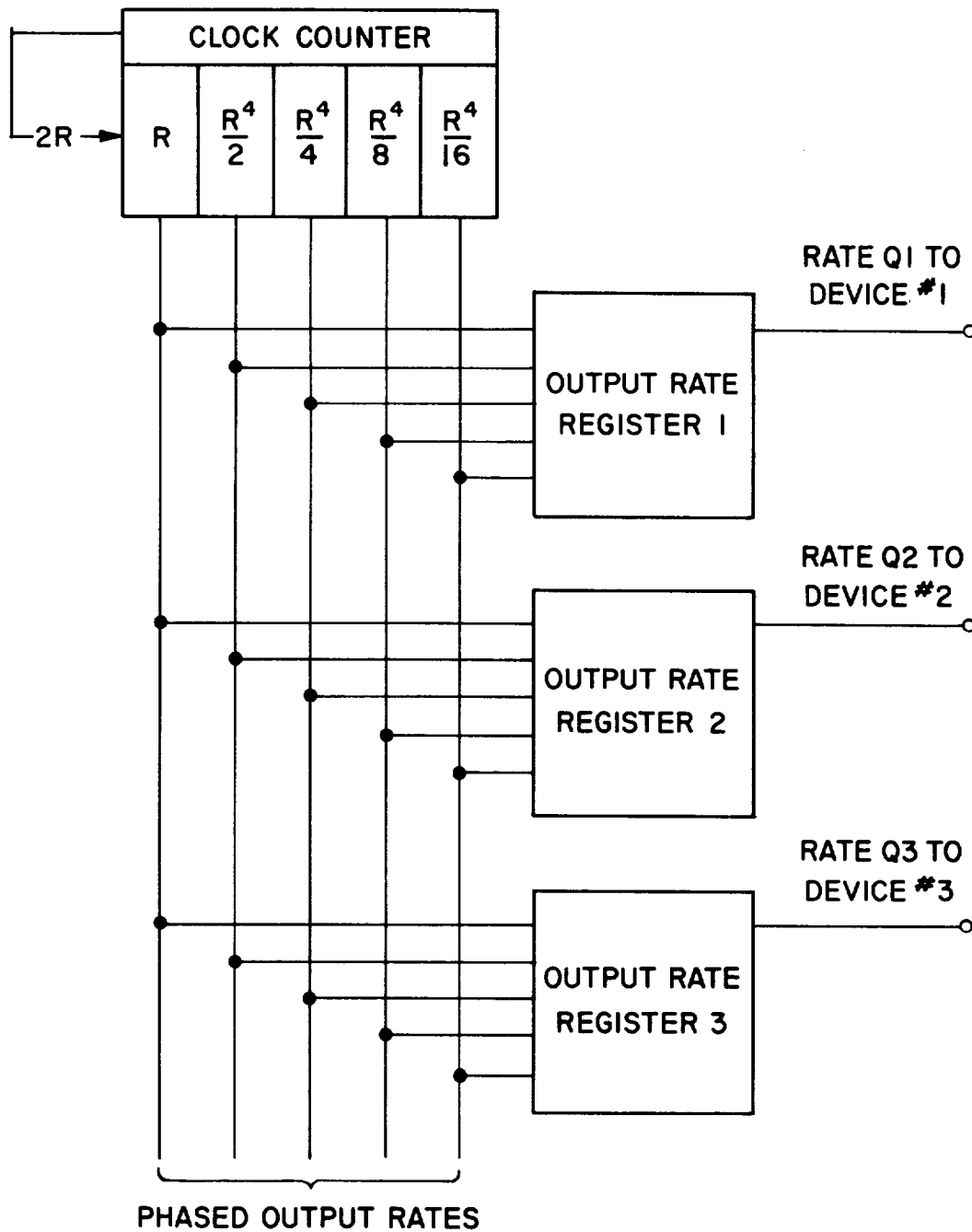


Fig. 3-11 Generator of several pulse rates. Pulse rate

$$Q = a_1 R + a_2 \frac{R}{2} + a_3 \frac{R}{4} + a_4 \frac{R}{8} + a_5 \frac{R}{16}, \text{ where}$$

$a_1 = 0$ or 1 , and R is some pulse rate determined by the clock. The contents of an output register determine which a are 0 and which are 1 .

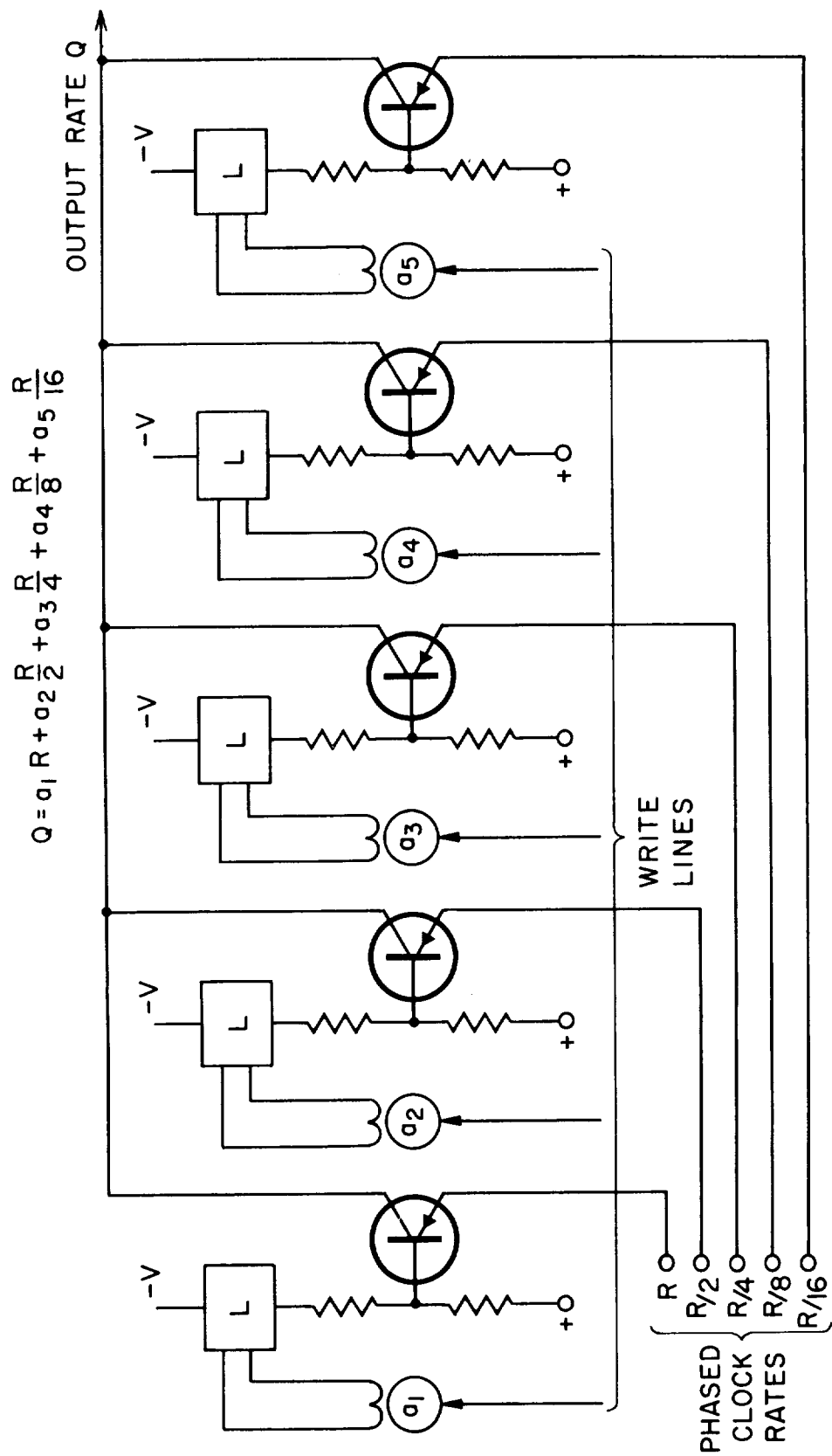


Fig. 3-12 Details of an output rate register.

where each a_i is either 0 or 1, and corresponds to a bit position. It is necessary^I to show that the various rates R , $R/2$, etc., can be generated in such a way that no pulses are lost due to time coincidence; for example, if the desired rate were $Q = 5R/8$, accomplished by combining rates $R/8$ and $R/2$, pulses would be lost if the wires carrying these respective rates had pulses on them at the same time.

The clock counter of Fig. 3-11 has an internal pulse source of frequency $2R$, and five flip-flops which successively halve the input rate. Thus, after every input pulse, the state of the Clock Counter can be expressed as a binary number. If the outputs R , $\frac{R}{2}$, $\frac{R}{4}$, etc., are taken to be the transitions of the halving flip-flops, as they change from state ONE to state ZERO, then there can be phasing problems, since it is possible to have several flip-flops going from ONE to ZERO simultaneously. However, there can be only one flip-flop going from ZERO to ONE at any one pulse time. In fact, there will always be a single flip-flop going from ZERO to ONE, with the exception of the pulse time in which the count goes from all ONES to all ZEROS. Therefore there are no phasing problems in the rates R , $\frac{R}{2}$, $\frac{R}{4}$, etc., are generated by the ZERO to ONE transitions of the halving flip-flops. The situation is illustrated in the tables of Fig. 3-13, which shows the generation of several output rates.

A variant of the Rate Generator discussed above which could be of some importance is the Rate Multiplier. The output Q of a Rate Generator is $Q = AR$, where A is the content of the Rate Generator. Thus, if R were a variable rate, Q would be the product of A and R . Rate Multipliers each require a set of counting flip-flops, so that they are fairly expensive in equipment*. However, all these special techniques are available for generating variable pulse rates independently of the main computer activity.

*For a discussion of this subject, see Grabbe, Ramo, Wooleridge, "Handbook of Automatic Controls and Computation".

Count	Clock Counter Output Pulses					Example 1	Example 2	Example 3
	$\frac{R}{16}$	$\frac{R}{8}$	$\frac{R}{4}$	$\frac{R}{2}$	$\frac{R}{1}$	$Q = \frac{5R}{16}$	$Q = \frac{3}{4}R$	$Q = \frac{31}{16}R \doteq 2R$
0 0 0 0 0								X ← Missing Pulse
0 0 0 0 1					1			1
0 0 0 1 0				1			1	1
0 0 0 1 1					1			1
0 0 1 0 0			1			1	1	1
0 0 1 0 1					1			1
0 0 1 1 0				1			1	1
0 0 1 1 1					1			1
0 1 0 0 0		1						1
0 1 0 0 1					1			1
0 1 0 1 0				1			1	1
0 1 0 1 1					1			1
0 1 1 0 0			1			1	1	1
0 1 1 0 1					1			1
0 1 1 1 0				1			1	1
0 1 1 1 1					1			1
1 0 0 0 0	1					1		1
1 0 0 0 1					1			1
1 0 0 1 0				1			1	1
1 0 0 1 1					1			1
1 0 1 0 0			1			1	1	1
1 0 1 0 1					1			1
1 0 1 1 0				1			1	1
1 0 1 1 1					1			1
1 1 0 0 0		1						1
1 1 0 0 1					1			1
1 1 0 1 0				1			1	1
1 1 0 1 1					1			1

Fig. 3-13 Pulse Rate Generation and Examples

CHAPTER 4

GENERAL REMARKS

A. Size, Weight, and Number of Components

In this section the size, weight and component quantity of the various subsystems presented before will be discussed.

1. The Core Rope

This subsystem is the most unusual one from the point of view of standard computer technology and packaging. The ropes constructed to date (Fig. 1-1) have a core density of 32 cores per in.³. Since each core represents one or two entire words of up to about 50 bits total, the bit density of the core rope could be about 1500 bits/in.³. This density does not take into account connectors, access equipment or mounting of the core rope on some fixture. As a better measure of size, a 4000 word memory, of 24 bits per word, will be postulated.

The core rope for this memory will consist of 2000 cores (two 24 bit words per core) and occupy 160 in.³, complete with mounting boards and end connectors. The forty-eight sensing amplifiers needed in this case require 25 in.³, and the S register and its drivers require 50 in.³. Hence the total volume of a 4000 word system would be 235 in.³, or 0.14 ft.³. It must be emphasized that this size is not a projection based on future improvements on existing packaging techniques, but a calculation based on an already constructed core rope of 256 cores. The unusually large bit density is achieved by the one core-per-word character of the memory, not by refinement of packaging. The

total weight of such a system, again including all connectors and mounting boards, would be of the order of 7 lb. Reasonable refinements on core rope packaging techniques could be expected to reduce both overall volume and weight by one third to one half the figures given above.

2. The Erasable Storage

A board containing 16 erasable storage registers, each of 12 bit capacity, has been constructed (Fig. 4-1). From the existing registers, the following calculations may be made for 24 bit registers. Each register, including the cores and transistors necessary for access to it, would measure $4.75 \times .5 \times .5 \text{ in.}^3$, or 1.2 in.^3 . The weight of such a register, not counting mounting boards and connectors, is about 20 gm. Including mounting board and end connectors, and assuming a board size which allows advantage to be taken of the register dimensions (the board of Fig. 4-1 does not), it should be possible to have a system of 128 erasable storage registers inside of 160 in.^3 , and which weighs about 5 lb.

3. The Central Registers

Except for the S, Z and Parity registers, all other central registers are electrically similar to ordinary erasable storage registers, and of comparable volume. The S register has already been counted in the discussion on the core rope. The Z and Parity registers require about three times the weight and volume of an ordinary register. Assuming a system of 10 central registers, plus Z and Parity; and making generous allowance for the volume required by control pulse circuitry; the central register complex should occupy a volume about as large as that occupied by 25 erasable storage registers, i. e., about 35 in.^3 , and weigh about 1.5 lb.

4. The Clock and Sequence Generator

These two items can fit a volume of less than 100 in.^3 , and weigh about 2 lb.

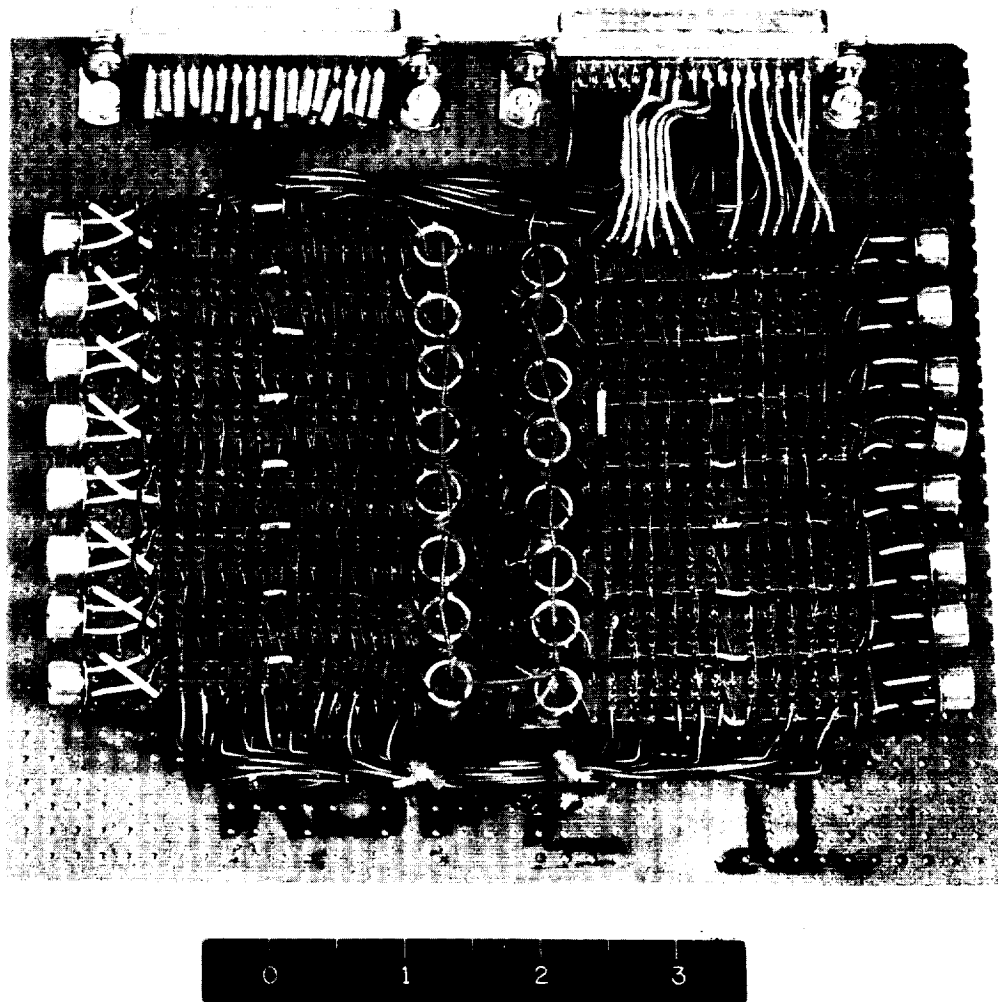


Fig. 4-1 16 word erasable storage system. (Clearing transistors at center. Write transistor-core combinations at edges. Write lines are under registers.)

5. Input and Output

An assignment of 100 in.³ and 4 lb to the input-output system seems reasonable to this writer. This would represent approximately 32 counter inputs (and counter registers), and 180 state inputs, figured at 6 state inputs in the volume occupied by two erasable storage registers.

6. Computer Totals

The many assumptions and calculations may be summed up as follows: a computer with a word length of 24 bits, with 4000 words of instructions and constants, with 128 erasable registers, and 32 counter and 180 state inputs, would occupy a volume of about 0.37 ft³, and weigh about 20 lb. It is interesting to note that these calculations give an overall density of 54 lb/ft³, which is approximately the density of other airborne digital computers. The various computer totals, including conservative estimates on the number of components, are summarized in Fig. 4-2.

B. Reliability

In discussing the potential reliability of the type of computer proposed here it is necessary to make several distinctions and assumptions. Primarily, it is necessary to distinguish between catastrophic failures e. g., burning out of a transistor, or opening of a solder joint, or breaking of a wire, and temporary failure due, for example, to voltage variations or a burst of extraneous radiation. It is assumed that a temporary failure does not result in any mechanical or electrical damage.

If there is sufficient program storage, and if speed of computation is not critical, then it is possible to protect against temporary failures by computing things in two different ways, at two different times, and checking results. A parity check system in which each word has a parity bit associated with it serves to detect the large majority of temporary failures and to initiate actions that call upon the computer's error correcting

	Volume in ³	Weight lb	Transistors and Latches made out of 2 Transistors	Diodes	Cores	Resistors	Comments
Rope Including S Register	200	6.0	60	—	2000	200	12 bit S register 5 T's/double bit
Sense Amplifiers and $\alpha \beta$ Amplifiers (48 bits)	35	1.0	480	—	—		10 T's/bit
Erasable Storage	160	5.0	256	3070	3070	25	128 registers, 24 bits each 2 T's/register
Central Registers and Control Pulse Amplifiers	35	1.5	168	300	240	40	12 central registers 60 control pulse amplifiers at 2T/amplifiers. 48 T's for P and Z
Clock and Sequence Generator	100	2.0	40	80	200	20	
Input - Output	100	4.0	450	1300	1300	400	
	630 in ³ .375 ft ³	19.5 lbs	1454	4750	6810	685	

Fig. 4-2 Physical characteristics of components.

programs.

Catastrophic failures are a more serious concern because of the difficulty in protecting against them. There are possibilities such as duplicating equipment, or having "spares" which can be connected automatically in place of a defective subsystem. These schemes are not attractive because of the expense in terms of volume and weight, and because it is not clear that they necessarily result in a more reliable system. The glib "connected automatically" may imply more reliability problems being generated than solved. The following paragraphs will contain a discussion of the likelihood of catastrophic failures in the proposed computer, without any duplications of equipment.

There is mounting evidence that a computer of the size and number of components of the one proposed here can be built so as to function for years without catastrophic failures. If it is assumed* that there will be no failures of a mechanical nature (poor construction and mounting), then the failures of concern are electrical ones. Of the various components listed in Fig. 4-2, cores may be eliminated as a source of electrical failures, for they may be damaged only by mechanical action or by temperatures above 300°C. The components most prone to failure are transistors and diodes.

The quality of transistors is improving steadily. The MIT Instrumentation Laboratory has had a computer in operation for over 7000 hours, in which there have been no failures among the 500 transistors used in the logic sections. Reports from Lincoln Laboratories** show similar results. Experience with diodes is similar to that with transistors. It is then reasonable

* This assumption does not imply a dismissal of the problem of construction as trivial.

** MIT-Lincoln Laboratories
Tech Report no. 199, ASTIA 214565, D. J. Eckl, P. A. Fergns,
R. L. Burke, "Transistor Life Experience", Nov. 1959.

to expect a system in which transistors and diodes numbering in the low thousands could be made so as to last for a few years without any catastrophic electrical failure. If the circuits are so designed that they are insensitive to changes in transistor parameters, the lack of catastrophic failures would imply no maintenance. The core-transistor circuitry which is the basis of the proposed computer is insensitive to changes in transistor β 's; the power dissipated in the transistors is well within the device's capabilities, and voltages are within breakdown limits.

The most serious remaining reliability problems are physical construction, sensitivity to temperature, and sensitivity to radiation.

Of these, the effects of radiation present in outer space are the least known.

C. Construction and Experience to Date

Various parts and circuits of the proposed machine have been built and tested in order to bring out logical and electrical problems, and to acquire some feeling for the relationship of the paper design to the actual physical system. This section will describe briefly what was built and how it worked. More detailed accounts will be published in future reports.

The most unusual element of the computer is its wired-in memory, and establishing the practicality of large ropes was felt to be a necessary first experiment. As a consequence, core rope systems of increasing sizes were built and operated. The largest system built to date has been one of 256 cores, and it has operated successfully. The basic electrical design problem consists of maximizing the transfer of energy from a pulse on the CLE line to the appropriate sense lines through a selected core, and minimizing the energy transferred through air coupling and through saturated cores. Air coupling was minimized relative to core coupling by essentially simple construction techniques, and by using as much iron in the cores as is practical with available

currents.

The sense lines were threaded through the rope cores in such a way that the noise generated by the CLE pulse through the cores that do not switch cancels out. The uniformity required of the cores is within commercial standards ($\pm 10\%$ variation in output signal for a given switching current). Sampling of the output signal was required only in the sense that the noises induced by the rise and fall of the inhibit currents are blanked out. The conclusion derived from the experience to date is that ropes of up to 4000 cores are perfectly feasible.

The erasable storage registers are of conservative design, and no serious problems have been encountered in that area. Various forms of pulse amplifiers, core-transistor pairs and shift registers have been built and tested; as a result there now exist definite design procedures which lead to successful circuits. A sequence generator has been built and operated, and also an 18 core priority circuit. The principles and circuits implied in the Z and P (Parity check) registers have been shown practical.

At this time an abbreviated version of the computer is being assembled, in which there are 32 erasable storage registers and 224 fixed registers (i. e., a 224 core rope). The word length is 12 bits, and there is a four instruction order code. This simple computer is expected to be useful in preliminary system tests.

The experience acquired to date indicates that the proposed computer is feasible and basically sound. The various design problems encountered do not call for refined circuitry or close component tolerances, and the components used are readily available commercially.

D. Power and Speed

It has been mentioned that with the proposed techniques a variable speed computer results in a correspondingly variable

power consumption. This feature is of obvious value in a space probe, where there is ample time to compute, and hence, where power may be saved. It is also of value in other airborne applications because it is possible to check out the computer at low speed, and not generate appreciable heat. In fact, it might be possible to keep the computer running continuously at very low speed and have a constant check on its functioning.

If the computer's oscillator clock should stop, no cores would switch, and the only currents would be those due to transistor leakage and those due to flip-flops^{*}. The leakage current of a transistor in the "off" condition is of the order of 10 μ a. If several transistors are in series, the leakage current is even less. For the representative computer, the total "stopped" current would not exceed 10 ma, exclusive of flip-flop current, which means 0.1 watt of standby power.

It makes sense to speak of the energy (in watt- μ sec) required for various operations such as clearing a register or switching a rope core. It turns out that it is possible to estimate the energy required by the actions resulting from the various possible control pulses. All control pulses except CLS and CLE require between 3 and 4 watt- μ sec. The inhibit currents generated by CLS have an amplitude of 300 ma, and last for about 5 μ sec, and are developed across a 10 volt power supply, so that they use 15 watt- μ sec. There are 12 inhibit currents, plus a set drive; hence the CLS pulse uses 185 watt- μ sec. Similarly, CLE uses 15 watt- μ sec. With these figures in hand it is further possible to estimate the energy of the various instructions. For example, the sequence for CA X has two occurrences of CLS and CLE, and 29 of other control pulses. On this basis, per-

^{*} Flip-flops are presently used in the sensing amplifiers and drivers of the prototype computer mentioned in Section 4C. Present plans call for the eventual replacement of these flip-flops by latch circuits which consume essentially no power in the "off" condition.

formance of CA X uses about 520 watt- μ sec. Energy required by other instructions is approximately the same. At full speed, i. e., performing an instruction every 48 μ sec, the estimated power consumption is approximately 11 watts. If flip-flops are used instead of latches in the Sense Write amplifiers (48 of them) then power consumption increases by about 4 watts. This increase is due to the fact that latches conduct, at most, during 3 μ sec out of every 12 μ sec clock cycle (the time between successive α pulses), whereas flip-flops conduct all the time at a rate of 10 ma. The total power consumption for the computer, running at maximum speed may then be estimated at 15 watts.